

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2001, 18:15:30 ; Search time 11120.8 Seconds
(without alignments)
3070.050 Million cell updates/sec

Title: US-09-284-180-4

Perfect score: 2315
Sequence: 1 ggggggtgcctctatgctgc.....cgttaataacatagtgctc 2315

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1283235 segs, 7373929652 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

GenBml:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_cm:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hcgo_hum:*
20: em_hcgo_inv:*
21: em_hcgo_rod:*
22: em_hcg_hum1:*
23: em_hcg_hum2:*
24: em_hcg_hum3:*
25: em_hcg_hum4:*
26: em_hcg_hum5:*
27: em_hcg_hum6:*
28: em_hcg_hum7:*
29: em_hcg_hum8:*
30: em_hcg_inv1:*
31: em_hcg_inv2:*
32: em_hcg_rod:*
33: em_hcg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_cm:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rod:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_v13:*
61: gb_v14:*
62: gb_v15:*
63: gb_v16:*
64: gb_v17:*
65: gb_v18:*
66: gb_v19:*
67: gb_v20:*
68: gb_v21:*
69: gb_v22:*
70: gb_v23:*
71: gb_v24:*
72: gb_v25:*
73: gb_v26:*
74: gb_v27:*
75: gb_v28:*
76: gb_v29:*
77: gb_v30:*
78: gb_v31:*
79: gb_v32:*
80: gb_v33:*
81: gb_v34:*
82: gb_v35:*
83: gb_v36:*
84: gb_v37:*
85: gb_v38:*
86: gb_v39:*
87: gb_v40:*
88: gb_v41:*
89: gb_v42:*
90: gb_v43:*
91: gb_v44:*
92: gb_v45:*
93: gb_v46:*
94: gb_v47:*
95: gb_v48:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2258	97.5	9 AX003081	AX003081 Sequence
2	2250	97.2	85 AB021292	AB021292 Homo sapi
3	2156	93.1	9 AX003083	AX003083 Sequence
4	1978	85.4	85 AB022317	AB022317 Homo sapi
5	1726	74.6	88 AF053369	AF053369 Homo sapi
6	1164	50.3	86 AC006543	AC006543 Homo sapi
7	1164	50.3	81 AL451003	AL451003 Homo sapi
8	1164	50.3	70 AC026541	AC026541 Homo sapi
9	1164	50.3	86 AC006544	AC006544 Homo sapi
10	1043	45.1	86 AC007387	AC007387 Homo sapi
11	623	26.9	9 AX003085	AX003085 Sequence

12	225	9.7	171964	81	AL451003
13	225	9.7	172010	70	AC026541
C	14	211	413	54	G22417
15	159	6.9	2669	85	AB047604
16	157	6.8	1581	88	AF119878
17	75	3.2	4008	94	AB002563
18	45	1.9	4002	94	AB021291
19	39	1.7	3205	94	AB022311S6
20	39	1.7	162651	94	AC003061
21	39	1.7	217260	60	AC007305
22	30	1.3	2649	94	AB022311S5
C	23	1.0	35080	5	CEL003H10
24	23	1.0	147788	84	AC015840
C	25	1.0	161039	83	AB003127
C	26	1.0	471188	85	AC004264
C	27	1.0	67332	90	AP001468
C	28	1.0	340000	90	AP001759
C	29	0.9	2637	1	AF170912
C	30	0.9	10658	63	AC013049
C	31	0.9	91658	60	AC005645
C	32	0.9	121332	60	AC007836
C	33	0.9	139027	61	AC010261
C	34	0.9	148113	81	AL391624
C	35	0.9	149607	67	AC022109
C	36	0.9	151687	80	AL356979
C	37	0.9	153280	82	AP001833
C	38	0.9	153946	60	AC008673
C	39	0.9	164318	76	AC079062
C	40	0.9	172127	65	AC019016
C	41	0.9	172670	65	AC018563
C	42	0.9	175826	67	AC021890
C	43	0.9	180653	80	AL359387
C	44	0.9	191736	69	AC025411
C	45	0.9	202402	63	AC012517
					AL451003 Homo sapi
					AC026541 Homo sapi
					G22417 human STR W
					AB047604 Macaca fap
					AF119878 Homo sapi
					AB002563 Rattus no
					AB021291 Mus muscu
					AB022316 Mus muscu
					AC003061 Mouse Chri
					AC007305 Mus muscu
					AB022313 Mus muscu
					U93832 Caenorhabdi
					AC015840 Homo sapi
					AP003127 Homo sapi
					AC004264 Homo sapi
					AP001468 Homo sapi
					AP001759 Homo sapi
					AF170912 Caulobact
					AC013049 Drosophill
					AC005645 Drosophill
					AC007836 Drosophill
					AC010261 Homo sapi
					AL391624 Homo sapi
					AC022109 Homo sapi
					AL356979 Homo sapi
					AP001833 Homo sapi
					AC008673 Homo sapi
					AC079062 Homo sapi
					AC019016 Homo sapi
					AC018563 Homo sapi
					AC021890 Homo sapi
					AL359387 Homo sapi
					AC025411 Homo sapi
					AC012517 Homo sapi

ALIGNMENTS

Db	626	GTGGGTGTGTCCGAGGACCTGGATTCCGACAGATACCTTGCCTCTTGCTGAACGCCCA	685
OY	127	gactctgtcgacagccgctgagccctttagccaccagccgaaatgggggagatgaagatgagagac	186
Db	686	GCTTTGTGCGAGCCGtgggCCTTtGAGCCAGCCGAATGGGGGATGAAGATGGACAGAC	745
OY	187	gaactcaactctctcttaccagagacttcccgaaattgactcalaogagcgatlaa	246
Db	746	GAATCTACTCTCTCTTtACGGAGACTCCCGAGCACTTGTAGCTCAATCACAGCGCATTTAA	805
OY	247	gtcccaaggtgtgcccgtgtgtgttgggggagacctggggggccgaagagccctccaag	306
Db	806	GTCCCAAGGGtggCCCTGTGTGTGGGGGAGCTTGggGGCCGGGAAGCCTTCCAGCAG	865
OY	307	agatgtacgacgcttcttgaagagcttgaacctgtctgtccagggccttgaatgtgccggcc	366
Db	866	AGATGTACGACGTTTtTGAAGCTGAACCTGTCTGTCCAGGGGCTGAGCATGGCCGGGC	925
OY	367	tccaggtctctgtcaggatgtgtgtgtcttgaaccttgaagcttggggcaggagctccatc	426
Db	926	TCCAGTGTCTTCGACAGGATtTGTGTGTCTTCGACCTGAGACCTTGGGGCAGGGACCTCCATC	985
OY	427	tttatatgacatctttcttccagatgggaaggagcacaatctctgcgtctgtctgtccctc	486
Db	986	TTTTATtGGACTTTTtTCTTCCAGtGGAGAGGGGCTACTATCTCTGCTGTCTGTCTTTC	1045
OY	487	cgacaccaagacattcggacagctgtcgtgaatgtgtcccttaagaagaaataaatactgac	546
Db	1046	CGACCAACAAGACTTCGAGCAGTGTGAATGGTCCCTTCAGAGAACTAAACAATCATCTGC	1105
OY	547	aacgagagactgtcctgtgcgtgagacaatgatgtgtccccaagccaagaccttggagatgtatc	606
Db	1106	AACGAGAGACTGCTCTGTGTGACAAATGATGTGCTCCCAAGCCAGACCTGGAGATGTATCT	1165

[illegible]

Db	1166	ACCAACACATGAAAGCTCCGGACATTGGCTATCTCTCCCTGCTACCGCGTACTC	12225
QY	667	accctacccggagcaaccaaccatcatatgaaagagccagtglttccagctatgtgccacc	726
Db	1226	ACCTTCATCCGGGACCCACCCACTCATGAGACAGGCCAGTGTTCAGCTGATGGCCACCCC	12851
QY	727	ctgtgtgttcactcaagataagcctctctcaagatcgatggccaaagggtgaacgcctc	786
Db	1286	CTGTGTGTCTACTACAGTATACAGCTTATCTCAGATGTGTGCCACAGGGTGACACGCTTC	13451
QY	787	tcaaggaaagagatagatgtatgtgtcctaacctggagacagagatgatgacaacctccacggagca	846
Db	1346	TCAGGAAAGAGTATGATGTGTCTTACCTGCGGGACAGAGAGATGGAACTCCACGGAGCA	14051
QY	847	gttcgcgatctggagctcagctcagcgcttcttgaagatctgtgcttatcccaagaccacag	906
Db	1406	GTCGGGATCCGAGAGCTACGCTACACGATCTGTGAAGATCTTGCCCTATATCCACAGCCACAG	14651
QY	907	ccagcttagagacatgaaattgtatcaacaagctgtgctcctgtgttgctcccgatactgaagtgtg	966
Db	1466	CCATTTAGAACATGAATATGTATGTAACACAGACTGGCTCTCTGTGGCTCCCGTACTAGGTG	15251
QY	967	acacaagtgaatacaacaacacgtgtgacgtctccagaagctgtctcagatgtcaatgtcaatccctg	1026
Db	1526	ACACAAATGTAAATACAAACCAACTGTGGCCGCTCTCCAGAGCTGCTCAGAGAGGCAATCCTGGCC	15851
QY	1027	caggagccaatctgtgtgctgtgagctctccggctcgtgatatgtgtgtgtgccaatgtgcgggag	10866
Db	1586	CAGAGACCACTGTGTGCTGTGAGACTTTCGCGCTGAGATGATGTGTGGCCCATCCCGGGAG	16451
QY	1087	caaccagagctgtgtcccaagacataagatcaagacatgtctccctcttctgtgtcctaaag	11461
Db	1646	CACCGAGGTTGGTCCAAAGACATAGAGTCAAGCAGATGTCTCTCTTTGTGTCTAAAGC	17051
QY	1147	cctgtgagaaacctcagatagtgttgaagttcccggtgtcacaagctgtgcatagtgtgtcttg	12061
Db	1706	CTGTGGAACAGTCTCAGTATAGTGTGAAGTTCCTGTGGCTACAGCTGCCATGTGGTCTTG	17651

QY 1207 ccaatgtctcaagctcagcattggcattcctgtgtgtgcaaccagcttgatgact 1266
|||||
DB 1766 CCATGTTCTCAAGCTCAGCATGGGCATCTGTGTGTGGCACCAGCCACTGAGTGA 1825
QY 1267 gcaatcaaccccccggcgagatgagcttgaggtaggttgaaccccaaggagcattggcgct 1326
|||||
DB 1826 GCACCTACCCCCGGGGGAGTGAAGTGTGTGACCCAGAGGGGCGCATGGGGCT 1885
QY 1327 tatgccttgtaattcagagaggttgaggcaccatgtgtatgacagcttcaagcttgta 1386
|||||
DB 1886 TATGCTGTGAATGTCAAGAGGGGTGGGGCCGCCCATGTGTGTACAGCTTACACTTGTA 1945
QY 1387 tggagcagcagcagagatgctccgagccggcccaacagtgaggcgaggacttgctgac 1446
|||||
DB 1946 TGGGCGACGCGAGCATGCTCCGAGCCGGCCACACAGTGGGGGGCGGAGCTGGCTGC 2005
QY 1447 tttcttttgggagatctcgaagatcccttgactctaatcttctgttggctggctcagag 1506
|||||
DB 2006 TTTCTTTGGGAGTCTCGACGATCCCTGACTCTCATTTCTGATTTGCTGGCTGACGAG 2065
QY 1507 cgaagcgagagagagagactctgctgctagagagaagtgaggcttgagcctggaggctcca 1566
|||||
DB 2066 CGACGGCGAGAGAGGAATCTTGCTGAGAGACAGAGTGGGCTTGAGCTGGGGGCTCCA 2125
QY 1567 cctctcggagcacaagctacagccaagaccctccctccctcctcgaagatgagcgg 1626
|||||
DB 2126 CTTTCTGGGACACAGACCTACACCAAGACCCCTCCCTCCCTGCTGTAAGATGAGCGG 2185
QY 1627 ttggcgttgagccttgagcagaaggggaggtgcttggatgattcacaacccctctg 1686
|||||
DB 2186 TTTCCCTGGGCTGCTGGCAGAGGGGAGTGGCTTGGTGGATTTCTACACCCCTTCTCG 2245
QY 1687 ctgagctctggcccaagcccaagccacattcggcttaactgggctccctcagccacatgt 1746
|||||
DB 2246 CTGTGATCTTGCCCAAGCCCAAGCCACATTCGGCTTAATCTGGGCTCTCTAGCCACATGT 2305
QY 1747 gatgaacatcattcattagagctgggcaaatgacacattgtatgaatgacttgagac 1806
|||||
DB 2306 GATGAACATCCTATCTAGACCTGGGCAATGACACATGATGATGATGATGATGATGATG 2365
QY 1807 ggggtgagccctgagagctgggggggtcactgggcttggaagacacccagctctgagc 1866
|||||
DB 2366 GGGGTGACCTAGAGTGTGGGGGTCTACTGGGCTGGAGACCATCCAGCCGTGTGAGT 2425
QY 1867 tctcttgagatgagtgatgattctgattctgattctctctctgagcctgagctg 1926
|||||
DB 2426 TCTCTTGAATGATGAGTATGATTACTTGTGATTTTGTCTCTCTGTGAGCCGTGATGG 2485
QY 1927 gcttggggcagaaccttgcctgattcctgattcccaatgaatcagaactgcttctg 1986
|||||
DB 2486 GCTTGGGGCGAGACCTTGGCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTG 2545
QY 1987 cagcaaatcagagcttcccccatacatcgaactcctgttaaacctcattccctggccccc 2046
|||||
DB 2546 CAGCAATATCGAGGCTTCCCTCAATCATCTGATCTCTGTAACCTTATCCCTGGCCCCC 2605
QY 2047 taacttgagccattagttcttgaggatgaggcacaagggcattagctatgacttgcttct 2106
|||||
DB 2606 TATCTTGGGCGCATTTAGTTTGGGGATGGGGCAGAGGCAATAGCTATGACTTTGCTTCT 2665
QY 2107 ggttggagccttgagcaggaagagccctggagggtggttggggggcacaatgtgcttga 2166
|||||
DB 2666 GGTGGGAGCTGGCGGAGAGAGAGAGCCCTGGAGGTGGGGGCAAAATGTGCCCTGAG 2725
QY 2167 tctcttgggggtggttctgcttctcattcaagttcattcgtatcgttgagggtgagatc 2226
|||||
DB 2726 TCTTGGGGGTGTTCTGCTTATTTCTCAAGTTTATCTGAATCTGTGGGGAGTGTGATATC 2785
QY 2227 ccaatgttgcaatagagctctgcccctgagatcttcccatctcagtttcccttccat 2286
|||||
DB 2786 CCATATGTGAATATGAGTCTGTGCCCTGAGATCTTCCCATCTCAGTTTCTTCCAT 2845

QY 2287 gaaagatcagtgtaatatcattgctc 2315
|||||
DB 2846 GAAAGATGACTGTGAATATCATGATGTTTC 2874

RESULT 2
AB021292
LOCUS AB021292 2672 bp mRNA PRI 09-APR-1999
DEFINITION Homo sapiens mRNA for semaphorin W, splicing variant deleting exons 5 and 6 and causing missense frameshift.
ACCESSION AB021292
VERSION AB021292.1 GI:4519449
KEYWORDS semaphorin W.
SOURCE Homo sapiens adult brain tissue_lhb:hippocampus cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Encinas,J.A., Kikuchi,K., Chedotal,A., de Castro,F., Goodman,C.S. and Kimura,T.
TITLE Cloning, expression, and genetic mapping of Sema W, a member of the semaphorin family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2491-2496 (1999)
MEDLINE 99162633
REFERENCE 2 (bases 1 to 2672)
AUTHORS Encinas,J.A., Ishida,H. and Kimura,T.
TITLE Direct Submission
JOURNAL Submitted (16-DEC-1998) to the DDBJ/EMBL/Genbank databases. Toru Kimura, Sumitomo Pharmaceutical's Research Center, 3-1-98 Kasugade-naka, Konohana, Osaka 554-0022, Japan (E-mail:tkimura@sumitomopharm.co.jp, Tel:81-6-6466-5228, Fax:81-6-6466-5491)
COMMENT Sequence updated (08-Jan-1999).
FEATURES
source location/Qualifiers
1..2672
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="adult"
/tissue_lhb="hippocampus"
/tissue_type="brain"
CDS
/note="splicing variant deleting exons 5 and 6 and causing missense frameshift."
/pseudo
/codon_start=1
/product="semaphorin W"
BASE COUNT 524 a 770 c 765 g 613 t
ORIGIN

Query Match 97.2%; Score 2250; DB 85; Length 2672;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 9999gtgtcctctatagtctgacactgtgaaaaaaccttgggagcaggaaccaattatcac 60
|||||
DB 372 GGGGCTGTCCTATAGCTGCACATGTGAATAACTACTGGGAGAGGCAATATATACCC 431
QY 61 agagcagtggtgtgtgcgcggagagctgattccggaacagatatcctgtcctcctgtgaac 120
|||||
DB 432 AGAGCATGGGTGCTGCGAGGAGTGGATTGCGACATACCTTGTCTGTGCTGTAAC 491
QY 121 gccccaagccttctgcagagcgttgacctgagcccaagcgaatgggggagtgaaatgga 180
|||||
DB 492 GCCCAGACCTTTGTGCGAGCGCTGCGCTTGTAGCCAGCCCAATGGGGGATGAAGATGGA 551
QY 181 gaagcgaatcactctctcttaagagagactcccgagcattgactcatacgaagcgc 240
|||||
DB 552 GAGGACGAATTAATCTTCTTTACGAGACATTCGCCGACATTTGACTATGAGACGCG 611
QY 241 attaagatcccaaggggtggccgctgtgtgtgctggggagagccttgggggcggaagaccctc 300
|||||
DB 612 ATTAAGTCCACAGGGGTGGCCGTGTGTGTGGGGGAGACCTCGGGGGCGGAAAGACCTC 671

OY 301 cagcagaatgagacgcttcttgaaagctgacgtctctgtccagggtcgtgaatgac 360
| | | | |
Db 672 CAGCAGAGATGACACGCTTTTGAAGCTGACCTGCTCTGCAAGGGCTGAGATGCG 731
| | | | |
OY 361 cggagctccagctgctgagagatgtgctgtctcgaacccgtgagcttgaggcaagagct 420
| | | | |
Db 732 CGGGCTCCAGTGTCTCTGAGAGATGCTGCTGTGCTTCAAGCTGAGCTTGGGAGGAGCT 791
| | | | |
OY 421 cccalccttaltgacatcttctctccagtgaggaggggtactactctctgtctgt 480
| | | | |
Db 792 CCCATCTTTTATGGCATCTTTCTTCCAGTGGAGGGGGCTACTATCTCTGCTGTCTGT 851
| | | | |
OY 481 gctcttcgagcccaagaacatctcgagcagtgctgaatgtgtccctcaagagaactaaacat 540
| | | | |
Db 852 GCTTTCGACCAAGACATTTGGACAGTGTGAATGGTCCCTTAGAAGAACTAAAGACT 911
| | | | |
OY 541 gactgcaagaagagagctgctgtgtgagcaatgagtgtcccaagcccaagacttgagag 600
| | | | |
Db 912 GACTGCACACAGAGAGAGCTGCTGTCTGACAAATGATGTGCCAGCCAGACTGGAGAG 971
| | | | |
OY 601 tgcaltcaaccaacaatgaagctccgacatlttgctcaltctctctccctgtccagcgc 660
| | | | |
Db 972 TGCATCACCAACAACATGATAAGCTCCGGGACTTTGGCTCATCTCTCTCCCTGCGTACCGC 1031
| | | | |
OY 661 gtaccacaccttaaccggagaccacacactcatggaagagccagtgcttccagctgaatgc 720
| | | | |
Db 1032 GTACTCACCTTCTATCCGGGACACCCACTCATGAGACGCCAGTGTTCACACTGATGGC 1091
| | | | |
OY 721 caacccctgtctgtactacagatacagactatcccaagctgtgagcccaagggtgac 780
| | | | |
Db 1092 CACCCCTGCTGTGTCTACTACATACAGCTTATCTCAGAGTGTGTGCCACAGGGTGAC 1151
| | | | |
OY 781 agcctctcaaggaaagatgatagtgtctctacccctggagacagaagatlgagacactccac 840
| | | | |
Db 1152 AGCCTCTCAGGGAAGAGATGATGTGCTCTACCTGGGACAGAGATGACACCTCCAC 1211
| | | | |
OY 841 cgagcagtgagagatctggagctcagctcagctgtcttgaaagatctgtgcttattcccaag 900
| | | | |
Db 1212 CGAGAGTGTGGATGGAGCTCAGCTCAGCGCTTGAAGATCTGGCTTATTCACAGG 1271
| | | | |
OY 901 ccacagcagctgtgagacatgaaatctgtacacagctgctcgtgtgtggtcccgact 960
| | | | |
Db 1272 CCACAGCCAGTTGAGAACATGAAATTGTACACACAGCTGCTCTGTGTGTGCTCCGCTACT 1331
| | | | |
OY 961 gagggtgacacaagtgaaatacaacacacacgtgtggcgctctccagagctgtcagagtgatc 1020
| | | | |
Db 1332 GAGGTGACACAAGTGAATACACCAACTGTGGCGCTCTCCAGAGCTGCTCAGAGTGCATC 1391
| | | | |
OY 1021 ctggcccaaggagacacagctgtgtgtcctgtgagctcccgctggatgtgtgtgcccatagc 1080
| | | | |
Db 1392 CTGGGCCAGGACCCAGTCTGTGCTGTGAACCTTCCGGCTGATGATGTGTGCCATATGCC 1451
| | | | |
OY 1081 gggagacacccggagctgtgtcccaagacatagatcagacagatgtctcctctgtgtcct 1140
| | | | |
Db 1452 GGGAGACACCCAGGGTGTGTCCAAACATAGAGTCAGCAGAGATGTCTCTCTTTGTGTCT 1511
| | | | |
OY 1141 aaagagccttgagaaagctccagtagtgcttgaaagctcccggtggctacagcttgagatgt 1200
| | | | |
Db 1512 AAAGAGCTTGGAAGAGTGCATAGTGTGAAGTTCCTGGTGGCTACAGCTCCGATGNG 1571
| | | | |
OY 1201 gtctgtgcatgtctcgaagctcagacatgggatactctgtgtgtgagacacagccagtgga 1260
| | | | |
Db 1572 GTCTTGCATGTTCTCAAGCTCAGCATGGGATCTCTGTGTGTGACACAGCCAGTGA 1631
| | | | |
OY 1261 gtgaactgacatcccccggcgggagatggaactggagtggtgtgtgtaaccccaaggggcatg 1320
| | | | |
Db 1632 GTGATGTGACATCAACCCCGGGGATGAGCTGAGAGTGTGTGTATACCCAGGGGCTATG 1691
| | | | |
OY 1321 ggcgtctatgctgtgaaatgtcagaagagtgaggagcccatgtgtgtagcagcttacagc 1380
| | | | |
Db 1692 GCGCTTATGCTGTGAATGTGACAGAGGTGGCGCAGCCATGTGTAGACACTTACAGC 1751
| | | | |

OY 1381 ttggtatggggcagccagcagatgctccagcgggcccacacagctggggcgaggctg 1440
| | | | |
Db 1752 TTGATATGGGGCAGCCAGAGGATGCTCCGAGCCCGGCCACACAGTGGGGCGGAGCTG 1811
| | | | |
OY 1441 gctggctctctcttgaggatctctgacagatccctgactctactctgattgtgtgaggt 1500
| | | | |
Db 1812 GCTGGCTTCTTTTGGGATTTCTCCAGATCTCTGACTCTCATCTGTGATGTGTGGGCT 1871
| | | | |
OY 1501 cagcagcagcggcagcagaggagactctgtgtctagaagaaggtgggacttgagc 1560
| | | | |
Db 1872 CAGCAGCGACGGGACAGAGGAGACTTCTGGCTAGAGACAAAGTGGGCTGGACCTGGGG 1931
| | | | |
OY 1561 gctccacactctgggacacaaagctacagcagaagacctccctccctctctctgaagat 1620
| | | | |
Db 1932 GCTCCACCTTCTGGGACCAACAAGCTACAGCCAAAGACCTCTCCCTCTCTCTAAAGAT 1991
| | | | |
OY 1621 gaggcgttgacagctgagccctgagcagaagaggagcaatgagctgtgtgtgtctccacaccc 1680
| | | | |
Db 1992 GAGCGGTGTCCCTGCTGGCTGGCCAAAGAGGGCACTGGCTTGTGTGATTTCTACCCACC 2051
| | | | |
OY 1681 ttcctgtgtatcctctgtcccaagcccaagcccaatctcgactgagctgggctctctaaagc 1740
| | | | |
Db 2052 TTCTGTCTGATCTTGGCCAAAGCCAGCCACATTCGGCTAATGGGGCTCTCTAGCC 2111
| | | | |
OY 1741 acatgtgatgaacatcactctagaagctggggaatgacactagtgataagtgatac 1800
| | | | |
Db 2112 ACATGTGATGAACATCCATCTAGAGCTGGGCAAAATGACCACTATGTATAGTATAC 2171
| | | | |
OY 1801 tggagcggagtgacacacagatgagatgtctgggggtcacttgagccctggaagacatccagcct 1860
| | | | |
Db 2172 TGGAAACGAGTACCACTGAGATGCTGGGGGTCACTGGGGCTGGAAAGACATCCAGCTT 2231
| | | | |
OY 1861 ctgagctctctcttgagatlgagatctacttgagatttagtactgtctctctgaagct 1920
| | | | |
Db 2232 CTGAGTCTCTTTGAGTATGAGATGATTAATGATTTAGTATCTGTCTCTGAGCCT 2291
| | | | |
OY 1921 ggaatgcttgaggccagacactctgtcctgatactccgatactccatgagaatcgaagctgc 1980
| | | | |
Db 2292 GGATGGGCTTGGGGCCAGACCTTGTGCTGATTCGATTCCTATGAAATCAGAACTGC 2351
| | | | |
OY 1981 ttctgcagcaaatcaaggagctcccccataacatctgaaactctgtaaaccttcaactctg 2040
| | | | |
Db 2352 TTTCTGCAAGCAAAATAGGGCTTCCCTTAACATCTGAACTCTGTAACTTTCACTCCCTG 2411
| | | | |
OY 2041 gccccctactcttgagcccatagatcttgaggatlggggacagagacatagctatgacttg 2100
| | | | |
Db 2412 GCCCCCTATCTTGGGCCCATTAATGTTTGGGATGGGCAAGGCATAGCTATGACTTGG 2471
| | | | |
OY 2101 cttctgtgtgagagctgtgcccgaagaaagagcccttgagatgtgtgtggggcaaatgtgc 2160
| | | | |
Db 2472 CTTGTGTGTGAGGCTGTGCGCGGAAGAGACCTCTGAGAGGTGTGGGGCAAAATGTGC 2531
| | | | |
OY 2161 cctgagctcttgagggtgtctgtctactctcgaagttatctgaaatctgtgtggaggtgc 2220
| | | | |
Db 2532 CCGTAGTCTCTTGGGGTGTGTGCTTATTCTTCAAGTTTATCTGAAATCTGTGGGAGTGC 2591
| | | | |
OY 2221 atgaltcccatgtgtgcaatlatgagatctctgacctgagatctctcccatctcaagttctc 2280
| | | | |
Db 2592 ATGATCCCATATGTCATATGAGATCTCTGCCGAGATCTCTCCCATCTCAGATTTC 2651
| | | | |
OY 2281 ttccatgaaagatagtgta 2301
| | | | |
Db 2652 TTCCATGAAAAGTACGTGTA 2672
| | | | |

RESULT 3
AX003083
LOCUS AX003083 2894 bp DNA
DEFINITION Sequence 3 from Patent EP0933425.
ACCESSION AX003083
VERSION AX003083.1 GI:9926965
KEYWORDS
SOURCE human.

ORGANISM	Homo sapiens
Euarchyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 2894)
AUTHORS	Michaelovich, D. and Doe, T.R.
TITLE	Semaphorin family polypeptides and polynucleotides
JOURNAL	Patent: EP 0933425-A 3 04-AUG-1999;
FEATURES	SMITHKLINE BEECHAM PLC (GB)
source	Location/Qualifiers 1..2894
BASE COUNT	/organism="Homo sapiens" /db_xref="taxon:9606"
ORIGIN	580 a 824 c 841 g 646 t 3 others
Query Match	93.1%; Score 2156; DB 9; Length 2894;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 2306; Conservative	0; Mismatches 3; Indels 0; Gaps 0
7	gtccctctatgctgcacactctgaaacacaccctgggagacgagccaattccaccagagca 66
566	gtccctctatgctgcacactctgaaacacactgacgggagacgagccaatttccacagagca 625
67	gtggagtcgtgcgcgagagactgatactcgagacagatacctctgcttcctgcgtgcgaagcccca 126
626	gtggagtcgtgcgcgagagactgatactcgagacagatacctctgcttcctgcgtgcgaagcccca 685
127	gaccttctgcagagccgttgagccttgagcccaagccgaattggggagatgagaatgagagagac 186
686	gcccttgcgcagccgttgagccttgagccttgagcccaagccgaattggggagatgagaatgagagac 745
187	gaactctact 246
746	gaactctact 805
247	gtcccaagggatggcccgctgctgctgcggggagacctcggggagccgggaagacccctccagcag 306
806	gtcccaagggatggcccgctgctgctgcggggagacctcggggagccgggaagacccctccagcag 865
307	agaatggagcagcttcttgaagctgaacctgctctctctctctctctctctctctctctctctct 366
866	agaatggagcagcttcttgaagctgaacctgctctctctctctctctctctctctctctctctct 925
367	tccagtgctctcgcagagatgttgtctgagcttgagccctgagccctggggagagagactccatc 426
926	tccagtgctctcgcagagatgttgtctgagcttgagccctgagccctggggagagagactccatc 985
427	tcttaatgcatcttctctctctccagctggagggggagctacacatctctgctgctgctc 486
986	tcttaatgcatcttctctctctccagctggagggggagctacacatctctgctgctgctc 1045
487	cgacccaagagacatctcgagcagtgctgtaattgctcccttcagagaacataacatgactgc 546
1046	cgacccaagagacatctcgagcagtgctgtaattgctcccttcagagaacataacatgactgc 1105
547	aacagagagactgcctctgctgtagaacaatgctgccccagccagagactggagagtgctc 606
1106	aacagagagactgcctctgctgtagaacaatgctgccccagccagagactggagagtgctc 1165
607	accacaacaatgaagctcccgagacttgctgatactctctctccctgctctgacccgctgactc 666
1166	accacaacaatgaagctcccgagacttgctgatactctctctccctgctctgacccgctgactc 1225
667	acctcatctcggaacacacccaacttcataagagagccagtgcttccagctgataagcagcccc 726
1226	acctcatctcggaacacacccaacttcataagagagccagtgcttccagctgataagcagcccc 1285
727	ctcgctgctacacagataagacactatctcagaatgctgagcccaagggctgagcaacccc 786
1286	ctcgctgctacacagataagacactatctcagaatgctgagcccaagggctgagcaacccc 1345
787	tcaagggaagagatataatgtgtctctcactctcgggagacagagatggaacactccacagagca 846

[illegible]

```

Db 2426 TCTCTTGTAGTATGATGATTTAGTATCTGTCTCTMTGAGCCTGATGCG 2485
Oy 1927 gcttgggagcagaccttgcctgatactcctgatacctgagaatacagaactgcttctg 1986
Db 2486 GCTTGGGCGCCACACCTTCTGATCTCTGATTTCCATGAGAAATCAGAACTGCTTCTG 2545
Oy 1987 cagcaatacagagcttcccccatacctcctgtaaacctcctgacccctgaccc 2046
Db 2546 CAGCAAAATCAGGCGTCTCCCTTACACTGAACTCTGTAACCTTCACTCCCTGCGCC 2605
Oy 2047 tacttgggagcattgatttgggagtgagcagagcatagatactgcttcttct 2106
Db 2606 TATCTTGGGCGCATTAAGTTTGGGAGTGGGCGCAGAGCATAGCTTGTGCTTCTTCT 2665
Oy 2107 gcttggagcctgagcaggaagagcctgagagtgagtgaggaggaactgagcctgag 2166
Db 2666 GCTTGGAGCCTGCGCGGAGAGAGCCCTCGAGGCTGTTGGGCGCAATGTGCCCTGAG 2725
Oy 2167 tcttgggagtgctgcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 2226
Db 2726 TCTTGGGCGTGTCTGCTTATCTTCAAGTTTATCTGAATCTGTGGGAGTGCATGATC 2785
Oy 2227 ccacatgctcaatagagctcctgagcctgagatcctccatcctcagcttcttctcat 2286
Db 2786 CCATGTTGCATATGAGATCTCTGCGCGAGATCTTCCCATCTCAGTTTCTTCTCAT 2845
Oy 2287 gaaagagtagctgtaatacatagcttctc 2315
Db 2846 GAAAGATGCTGTAAATACATAGTCTTCTC 2874

RESULT 4
AB022317 2400 bp mRNA PRI 09-APR-1999
LOCUS Homo sapiens mRNA for semaphorin W, complete cds, spliced variant
DEFINITION missing exons 5 to 9.
ACCESSION AB022317
VERSION AB022317.1 GI:4519502
KEYWORDS semaphorin W.
SOURCE Homo sapiens adult brain, hippocampus cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Encinas,J.A., Kikuchi,K., Chedotal,A., de Castro,F., Goodman,C.S.
and Kimura,T.
TITLE Cloning, expression, and genetic mapping of Sema W, a member of the
semaphorin family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (5), 2491-2496 (1999)
MEDLINE 99162633
REFERENCE 2 (bases 1 to 2400)
AUTHORS Jeffrey,E.A., Ishida,H. and Kimura,T.
TITLE Direct Submision
JOURNAL Submitted (11-JAN-1999) to the DDBJ/EMBL/GenBank databases. Toru
Kimura, Sumitomo Pharmaceuticals Co. Ltd., Research Center, 3-1-98,
Kaasugade-naka, Konohana, Osaka 554-0022, Japan
(E-mail:tkimura@sumitomopharm.co.jp, Tel:81-6-466-5228,
Fax:81-6-466-5491)
FEATURES
source
1..2400
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/dev_stage="adult"
/map="2p13"
/lisse_type="brain, hippocampus"
/gene="Sema W"
/gene="Sema W"
16..1863
/gene="Sema W"
16..1863
/gene="Sema W"
/note="spliced variant missing exons 5 to 9"
/codon_start=1
/product="semaphorin W"

```

```

/protein_id="BA75631.1"
/db_xref="GI:4519503"
/translation="MPASARPRPCPGPTASPELLLLAVLSGVSGRVSRVPSRTS
LPISADSCLEFRFAVPHYNYSVLLVDASHTLVYAGDTLFAISLPSGSRPRIDW
MPEARHONCKRKKKEDGLGRKTLQORHMTPLKADLCGPEHRSAYLQDVAVL
RPELCAGPRIFYGIFSSOWEGATISAVCAFPPDRIYLVNPPFRRLKDKRGLPYVD
NDVDPKRGECITNNMKLRHFGSSLSLSDRLVLTFRDPLMDRFPVFDGHPVLVTD
TAYLRVAHRYTSLSGKEYDVLVLTDEGHLHRAVRAIGQLSVLEDLALFPPQPVEN
MKLYHSMILVSRTEVQVNTNGRSLQSCSECLADPPVAMNFRDECVAAHGEHR
KLVDIESADVSSLCPEKPEGRPVVFEVPAVTAHVLPSPSSAMASCVHOPSGVT
ALTPRQDLEVVYVPMGAMVACQEGEAGAAHVVAASLVWGSORDAPSRATYAGIL
AGFGLITLAASTLILIGRROORPELLARDDVYGLDVGDPSPSTYSODPPSPSP
EDERPLALARGSGFSGFPPELDPSPFAHRLRGLATCDENST"
BASE COUNT 467 a 697 c 683 g 553 t
ORIGIN
Query Match 85.4%; Score 1978; DB 85; Length 2400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 273 ggggagcctcgggggagcagaccctccagagagatgagcagcttttgaagctga 332
Db 372 gggggagcctcgggggggggagagaccctccagagagatgagcagcttttgaagctga 431
Oy 333 ccttctctcagagcctcagatgagcagcctccagctgctcctcagagatgctgct 392
Db 432 CCTGCTCTGCCAGGCGCTGAGCATGCGCGGCGCTCCACTGCTGAGAGATGCTGCT 491
Oy 393 gcttggagcctgagcctgagcagagagatcctccttcttcttcttcttcttcttcttct 452
Db 492 GCTTGCACCTGAGCTTGGGCGAGGAGCTCCACTTTATGCGACTCTTCTTCCAGTG 551
Oy 453 ggaaggagctatctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 512
Db 552 GGAGGGGGCTACTATCTCGTGTGCTGCTTCCGACGACAAAGCATTTGGGAGAGTGT 611
Oy 513 gaatggtcctcctcagagaaataaatactgacagcagagagctgctcgtgagaa 572
Db 612 GATGTCTCTTCCAGAGACTAAACATGATGCAACAGAGAGCTGCTGCTGAGACAA 671
Oy 573 tgatgtgcccagccagagcctgagagatgacaaacaaacaaatgaagctcagcagct 632
Db 672 TGAATGTCCTCCAGCCAGCTGAGAGTGCATCACCAACATGATGCTCCGCGACTT 731
Oy 633 tggctatctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 692
Db 732 TGGCTCATCTCTCCCTGCTGAGCGGCTACTCACTTCACTCCGGACCCACTCAT 791
Oy 693 ggaagagcagctgctcagctgagtgagcagcctcctcctcctcctcctcctcctcctcct 752
Db 792 GGACAGGCGCAGTGTCTCCAGCTGATGCGCCACCCCTGCTGCTGCTACAGATACAGCTTA 851
Oy 753 tctcagagctcgtgcccagaggtgacagcctcctcagagagagagatgagtgctcctc 812
Db 852 TCTCAGAGCTGCGCCAGAGGCTGACAGCTCTCAGGGAAGATGATGATGCTCTCA 911
Oy 813 ccttggagcagagagatgagcagcctcagcagcagctgagatcgagagctcagcagct 872
Db 912 CCTGGGAGAGAGATGAGACACTCCACGAGACAGTGGCGATGAGAGCTCAGCTCAGCT 971
Oy 873 tcttgaagatctgacctatcctcagagagcagcagcagcttgaagaaatgaaatgta 932
Db 972 TCTTGAAGATCTGCGCTTATCTCCAGAGCCACAGCAGCTTGAACATGAAATTTGACCA 1031
Oy 933 cagctgagctcctgctgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 992
Db 1032 CAGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1091
Oy 993 ccgtctccagagctcagagtgatccttggccagagcagcagctgctgagcagct 1052
Db 1092 CCGTCTCCAGAGCTGCTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1151

```

[illegible]

Db	2232	CCCTGAGAGTGGATTGGGGGCAAAATGATGCCCTGAGTCCTGGGGGTGGTCTGCTAATTTCT	2291
Qy	2193	caagttatctaaactctgagggagtgatcatgtatcccaatgttgacaataatgagctctgc	2252
Db	2292	CAAGTTATTCATGAACTCTGTGGGGAGATGCATGATGCCCATGTTTGCAATATGAGACTTCTGC	2351
Qy	2253	cctgagatcttcccatctcagttctcttcctcatgaagaagatcgtga 2301	
Db	2352	CCTGAGATCTTCCCATCTGCACTTTCTTCCTCATGAAGAAGACTGTGA 2400	
RESULT	5		
LOCUS	AF053369	1913 bp	PRI 02-JUL-2000
DEFINITION	Homo sapiens semaphorin M (SEMAM) mRNA, partial cds.		
ACCESSION	AF053369		
VERSION	AF053369.1	GI:8894933	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1913)		
TITLE	Jang,W., Spilison,S.V., Hua,A., Roe,B. and Weisler,M.H.		
JOURNAL	large-scale comparative sequence analysis of human and mouse genomic DNA in the md2 region of mouse chromosome 6 reveals coding regions of three new genes		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1913)		
TITLE	Jang,W., Spilison,S.V., Hua,A., Roe,B. and Weisler,M.H.		
JOURNAL	Submitted (11-MAR-1998) Human Genetics, University of Michigan, 2806 Medical Science II, Box 0618, Ann Arbor, MI 48109-0618, USA		
FEATURES	Location/Qualifiers		
Source	1..1913		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/db_xref="dbEST:H10623"		
	/clone="IMAGE 47298"		
	<1..1913		
	/gene="SEMAM"		
	<1..1277		
	/gene="SEMAM"		
	/codon_start=3		
	/product="semaphorin M"		
	/protein_id="AA80660.1"		
	/db_xref="GI:8894934"		
	/translation="PODIRTVNLNPFRELKHDNCNRLPVVDNDVPOPRGECITNNMKMLHNSGELSLPDLVTRFDHMDRDPVDFPGLVTTTATLRYAVHRTSLSGKAEVDLYLSTEDGHLHRAVRIGAGOLSYLEDLALPEPQVEMKTLKSHLLVSTPYTQVNTNCRLOSCECLLADOPVCAMSFRDECYAHKEHGLVODIESAVSCLPKGEPEVPEVEVPAVTAHVHVLPCSPSSAMASCVMHOPSGLVTPRRDGLLEVVTTPAAGAACDECGGAHVVAAYSLVWGSQRDAPSRAHVTGAGIACGFLGILAASTLILILIGRQORRORRELLARDKVDGLDLCAPSGSTTSYSDPDSPSPEDERLPLALAKRSGCGFGSPFLPDPPSPAHIRLITGAPLATCDETSI"		
BASE COUNT	407 a 521 c 529 g 438 t	18 others	
ORIGIN			
Query Match	74.6%; Score 1726; DB 88; Length 1913;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1826:	Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
Qy	488	gaccacaagacatctcgagacatgctgaatgcttccttcagagaaactaaacatgactgca	547
Db	1	GACCACAAGACATTCGCGACACGCTGGAATGCTCCCTTCAGAGAATAAACAATGACTGCA	60
Qy	548	acaagagactgctctgtgtgacaaatgattgccccagcccgagactgtagagtgatcatca	607
Db	61	ACAGAGGAGATGGCTCTCGTGGAACAATGATGTGCCCAAGCCAGACCTGGAGAGTGCATCA	120
Qy	608	ccaacaacatgtagctcgcgcaacttgctgctacatctctctccctgctgacacgcgtactca	667

Db 121 CCAACACATGAAAGCTCCGGCACTTGGCTCATCTCTCCCTGACCGGCTACTCA 180
 Qy ccttcaaccgggaccacccactcaatgagagggagtgattccagctgaatggacacccc 727
 Db 181 CCTTCATCCGGGACACCCACTCATGAGAGGCGAGTGTTCAGAGCTGATGGCCACCCC 240
 Qy 728 tgcgtgcaatcaatgataagctatacctcaagatcgatggccacaggtgacacgctct 787
 Db 241 TCTCTGCTCACTAGATGATACAGCTTATCTAGATCTGTGGCCACAGGATGACAGCCTCT 300
 Qy 788 caggagaaagatgataatgctgctcaccctgggacagagatgagacccctccacagagac 847
 Db 301 CAGGAAAGATGATGATGCTCTTACCTGGGGACAGAGATGAGACCTTCACCGAGCAG 360
 Qy 848 tgcgagctgagagctcagctcagctctctgagagatctggccttatccagagacagag 907
 Db 361 TCGGATCGAGAGCTCAGCTCAGGCTTCTTGAAGATCTGGGCTTATTCAGAGCAGCAG 420
 Qy 908 cagctgagagacatgataatgatacagagctgctcctgctgctgctcctgctgctgctg 967
 Db 421 CAGTTGAGACATGAAATGTTACACAGCTGGCTCTGTTGGCTCCGCTACTGAGAGTGA 480
 Qy 968 ccaagatgataaacaacactgagcctctcagagagctgctcagagatgacatccctgagcc 1027
 Db 481 CACAGTGAATACACCAACTGTGGCGCTCTCCAGAGCTGCTAGAGTGCATCTGAGCC 540
 Qy 1028 agagaccagctctgctgctgagagctcagagctgagatgctgctgctgctgctgctgctg 1087
 Db 541 AGGACCACTGTGCTGCTGAGAGCTTCGGGCTGATGATGTTGGCCATGCGGGGAG 600
 Qy 1088 accgagagctgctgctgagagacatagagctgagagatgctcctctctgctgctgctgctg 1147
 Db 601 ACCGAGGCTGTGCTGCTGAGACATAGAGTACAGACATGCTCTCTTGTCTCCAAAGAG 660
 Qy 1148 caggagaaagctcagagatgctgctgagagctcagagctgagatgctgctgctgctgctg 1207
 Db 661 CTGGAGAACGCTCAGTAGTGTGTTGAAGTTCCTGGCTGACAGCTGCGCATGTGCTTGC 720
 Qy 1208 catgtctcacaagctcagagatgagacatcctgctgctgctgctgctgctgctgctgctg 1267
 Db 721 CATGTTCTCAAGCTCAGCATGGGCTCTCTGTGTGACACAGCCAGTGGAGTACATG 780
 Qy 1268 caatcaaccccccggggagctgagctgagagtggtggtgagacccagagggccatggcgctt 1327
 Db 781 CACTCACCCTCCCGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 840
 Qy 1328 atgagctgagatgctgagagagtggtggtgagagccatgctgctgctgctgctgctgctg 1387
 Db 841 ATGCTGTGATGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 Qy 1388 ggggagcagcagagatgctcagagctgagagccagacagtggtggtgagagagctgctgctg 1447
 Db 901 GGGGAGCAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 960
 Qy 1448 tcttcttgggagctcagagagctcagagctcagagctcagagctcagagctcagagctcag 1507
 Db 961 TCTTCTTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1020
 Qy 1508 gacgagcagagagagagctcagagctgagagagagtggtggtgagagagctgctgctgctg 1567
 Db 1021 GACGGGAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1080
 Qy 1568 cctctgagagcagagctcagagcagagacccctcctcctcctcctcctcctcctcctcctg 1627
 Db 1081 CTTTGTGAGCAGCAGAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1140
 Qy 1628 tggcgctgagcctgagcag 1687
 Db 1141 TCCCTGTGCTGCTGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
 Qy 1688 tgcagcttggcagagcagcagcagacatcagctgagagagagagagagagagagagagag 1747

Db 1201 TTGATCTTGCCCAACGCCACATTCGGCTAACTGGGGCTCTCTAAGCCATATG 1260
 Qy 1748 atgaaacatccatctagagctgggcaaatgagacacatgataagtgatacctgagag 1807
 Db 1261 ATGAACATTCATCTAGAGCTGGGCAAAATGACACATAGTATAGTATACATGGAAG 1320
 Qy 1808 gagtgacacatgagatgctgggggtgctacgtggcctgagaaacacacacacacacacac 1867
 Db 1321 GAGTGACCACTGAGATGCTGGGGGTGCTGAGGCTGGGAGACCATCCAGCCCTGAGT 1380
 Qy 1868 cctctgagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1927
 Db 1381 CTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Qy 1928 cttgggagagaccttgcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1987
 Db 1441 CTGGGGCAGACCTTGGCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 1500
 Qy 1988 agcaatcagagcttccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2047
 Db 1501 AGCAATCAAGGCTTCCTCAATCTGATCAATCTGATCAATCTGATCAATCTGATCAAT 1560
 Qy 2048 atctgggagcagatgcttgggagatgggagagagagagagagagagagagagagagag 2107
 Db 1561 ATCTTGGCCCATTAATTTTGGGATGGGACAGGACATAGCTATGATCTTGTCTG 1620
 Qy 2108 gttgagccttggcggag 2167
 Db 1621 GTTGGAGCCCTGGCGGAGAGAAANACCTGAGAGCTGTTGGGGCAAAATGTCCTGAGT 1680
 Qy 2168 ccttggggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 2227
 Db 1681 CTTTGGGGTGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
 Qy 2228 ccatgttgcataatgagatgctgctgctgctgctgctgctgctgctgctgctgctgctg 2287
 Db 1741 CCAATTTGCAATGAGAGCTCTGCTGAGATCTTCCCATCTCAGTTTCTTCTGATG 1800
 Qy 2288 aaagagtagtgataatagatgctgctgctgctgctgctgctgctgctgctgctgctg 2315
 Db 1801 AAAGACTAGCTGTAATATACATAGTCTC 1828
 RESULT 6
 AC006543 LOCUS PRI 27-MAY-2000
 DEFINITION Homo sapiens chromosome 2p13 clone bac172k1, complete sequence.
 AC006543.7 GI:4388738
 VERSION
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM
 human.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 1 (bases 1 to 126228)
 Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
 Homo sapiens Chromosome 2p13 BAC Clone bac172k1
 Unpublished
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 2 (bases 1 to 126228)
 Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
 Direct Submission
 Submitted (09-FEB-1999) Department of Chemistry and Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 3 (bases 1 to 126228)
 Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
 Direct Submission
 Submitted (09-MAR-1999) Department of Chemistry and Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 4 (bases 1 to 126228)
 Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
 Direct Submission


```
coverage: 4.10x in Q20 bases; agarose-1p
-----
Draft Sequence Produced by Whitehead Institute/MIT Center for
Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA
http://www-seq.wi.mit.edu.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 3477: contig of 3477 bp in length
* 3478 3577: gap of 100 bp
* 3578 13586: contig of 10009 bp in length
* 13587 13686: gap of 100 bp
* 13687 18688: contig of 5002 bp in length
* 18689 18788: gap of 100 bp
* 18789 26165: contig of 7377 bp in length
* 26166 26265: gap of 100 bp
* 26266 52879: contig of 26614 bp in length
* 52880 52979: gap of 100 bp
* 52980 82107: contig of 29128 bp in length
* 82108 82207: gap of 100 bp
* 82208 88054: contig of 5847 bp in length
* 88055 88154: gap of 100 bp
* 88155 91774: contig of 3620 bp in length
* 91775 91874: gap of 100 bp
* 91875 101902: contig of 10028 bp in length
* 101903 102002: gap of 100 bp
* 102003 109337: contig of 7335 bp in length
* 109338 109437: gap of 100 bp
* 109438 113399: contig of 3962 bp in length
* 113400 113499: gap of 100 bp
* 113500 118719: contig of 5220 bp in length
* 118720 118819: gap of 100 bp
* 118820 123815: contig of 4996 bp in length
* 123816 123915: gap of 100 bp
* 123916 132492: contig of 8577 bp in length
* 132493 132592: gap of 100 bp
* 132593 137709: contig of 5117 bp in length
* 137710 137809: gap of 100 bp
* 137810 149420: contig of 11611 bp in length
* 149421 149520: gap of 100 bp
* 149521 156195: contig of 6675 bp in length
* 156196 156295: gap of 100 bp
* 156296 171964: contig of 15669 bp in length.
FEATURES
source
1. 171964
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-136B23"
/clone_lib="RPC1-11.1"
1. 3477
/contig="assembly-fragment:00781"
clone_end:SP6
vector_side:left"
misc_feature
3578..13586
/contig="assembly-fragment:00011"
misc_feature
13687..18688
/contig="assembly-fragment:00077"
misc_feature
18789..26165
/contig="assembly-fragment:00151"
misc_feature
26266..52879
/contig="assembly-fragment:00273"
misc_feature
52980..82107
/contig="assembly-fragment:00398"
misc_feature
82208..88054
/contig="assembly-fragment:00431"
misc_feature
88155..91774
```

```
misc_feature
/contig="assembly-fragment:00477"
91875..101902
/contig="assembly-fragment:00705"
misc_feature
102003..109337
/contig="assembly-fragment:00744"
misc_feature
109438..113399
/contig="assembly-fragment:00927"
misc_feature
113500..118719
/contig="assembly-fragment:00984"
misc_feature
118820..123815
/contig="assembly-fragment:00995"
misc_feature
123916..132492
/contig="assembly-fragment:01012"
misc_feature
132593..137709
/contig="assembly-fragment:01113"
misc_feature
137810..149420
/contig="assembly-fragment:01399"
misc_feature
149521..156195
/contig="assembly-fragment:01504"
misc_feature
156296..171964
/contig="assembly-fragment:00674"
clone_end:77
vector_side:right"
BASE COUNT 48050 a 38057 c 35374 g 48570 t 1713 others
ORIGIN
Query Match 50.3% Score 1164; DB 81; Length 171964;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1152 aagaacgtccagtagtcttgaagatccctggtcagcagctgcatgcttggtcagcagc 1211
|||||
DB 80003 AGAAGCTCCAGTAGTGTGTAAGTTCCTCGGTGATACAGTCCGACATGCTCTGCCATG 79944
QY 1212 ttctcaagctcaagctggtgcatctctgtgtgtgcaacgacccagtggtgagtcgact 1271
|||||
DB 79943 TTCTCCAACTCAGCATGGGATCTCTGTGTGTGGCACGACCCAGTGGAGTCACTGCACCT 79884
QY 1272 caaccccgcgagatgagctgaggtgtgtgtgtgacccagggcgcatggtcattgc 1331
|||||
DB 79883 CACCCCGCGGAGATGAGTGGAGTGGTGTGTGACCCAGGGGCGATGGCGCTTATGC 79824
QY 1332 ctgtgaatgtcaggaaggtgtggtgacgacatgtgtgtgtagcagctttagttagtgg 1391
|||||
DB 79823 CTGTGAATGTCAAGAGAGGTGGGGACGCCATGHTGATGACCTTGTGATGGGG 79764
QY 1392 cagccagcgagatgtctcgaagccgggcccacacagctgtggtggtcggtgtcttct 1451
|||||
DB 79763 CAGCCAGCGAGATGCTCCAGCGGGGCCACACAGTGGGGGGGAGCTGTGGCTTCTT 79704
QY 1452 ctgtggaatctcgcagagatccctgactcattcgtatgtgtgtgtgtgtgtgtgtgtgt 1511
|||||
DB 79703 CTGTGGGATTTCTCGCAGCATCTCTGACTCTCATTTCTGATTTGTGGCGTCAGAGGAGG 79644
QY 1512 gcgacagaggaactctgtcctagagacaaggltggtgctgtgacctgtggtggtgtgtgt 1571
|||||
DB 79643 GCGACAGAGGAACTCTGTGCTAGAGACAAGGTGGGCGCTGTGACCTGTGGGCGTCCACTTTC 79584
QY 1572 tgggacacaaagctacagcacaagaccctccctccctcctcctgaagatgaagcgtttgc 1631
|||||
DB 79583 TGGGACACAAAGCTACAGCAAGACCCCTCTCCCTCTCTGAAGATGAGCGGTTGCC 79524
QY 1632 gcttgccctgtgccaagaggtgagtgcttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1691
|||||
DB 79523 GCTTGCCCTGTGCCAAGAGGGAGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 79464
QY 1692 tccctgtcccaagcccaagccacatctcgtctaaactgtggtgtgtgtgtgtgtgtgtgtgt 1751
|||||
DB 79463 TCCCTTGCCCAAGCCCAAGCCCATTCGCTTAACGTGGGCTCCTGTAGCCACATGTGATGA 79404
QY 1752 aacatccatctagagctgtggtgcaaatgagcactagtgatataagatgatactgtgaagga 1811
|||||
```

Db	79403	AACATTCACATCTAGAGCTGGGCAAAATGACACACATAGTGTATTAAGTATGATCACTGGAAGGAGGT	79344
Qy	1812	gaacacatgaagatgctcggggggtcaatcggccctggaagaccatcccaagcctctgaatctctc	1871
Db	79343	GAACCACTGAGATGCTGGGGGGTCACTGGGGCTTGGAAGACCATCCAGCCCTGAGTTCTCT	79284
Qy	1872	ctgaatctaaagagatacttgaaatttgtagtctgtctctctctgagccctggaatggcttg	1931
Db	79283	TTGACATATGATGATTAAGTACTGGATTTTAAATATCTGTCTCTCTGAGCCGTGATGGGCTTGG	79224
Qy	1932	gggccaagaccctctgcctgcatctccgatcccaatggaatccagaactgcctctctcaaga	1991
Db	79223	GGGCGACAGACTTTGGCTGATTTCTCTATTCCTCCATGGAATAATCGAAGCTGTTTCTGACGA	79164
Qy	1992	aatcagaagcctcccccataacatctgaactccctgtaaacctccatcccttgagcccccatact	2051
Db	79163	AATCAGGGGCTTCCCCCTTAACATCTGACATCTGTAACCTCTGTAACCTCATCCGTGGCCCCATCT	79104
Qy	2052	tggagcccatagatcttttgggagatgggggacaaagggacatagctatgaacttggcttctggtg	2111
Db	79103	TGGGCCCAATTAATTATTGGGAGTGGGGCACAGGGCATAGCTAATGACTTTGCTTTCGGTTGG	79044
Qy	2112	gaagcctggaccggaaagaaagaagccctcgaggggctggtcgggggcaaaatgctgcctcgagccctt	2171
Db	79043	GAGCCTTGCCGGGAAAGAGAAAGCCCTGGAGAGTGTGGGGGCAAAATGTGCCCTGAGCCCTT	78984
Qy	2172	gggggtgctctcgtcatctctcaagtcttaacgtcttaacgtacatcgtcggagagtgcatatcccat	2231
Db	78983	GGGGGTGTTCTCTCTATTCTTTCGAAGTTTATCTGAAATCTGTGGGGAGTGAATATCCCCAT	78924
Qy	2232	gttgcaataatgagatctctcgcctctgagatcctcccatcctcaagtcttccctccatgaagaag	2291
Db	78923	GTTGCAATATGAGATGCTCTGCCCCGAGATCTTCCCATCTCAGTTTCTCTCCATGAAGAAG	78864
Qy	2292	agtaagctgtaaatacatagtgctc 2315	
Db	78863	ACTACGTGTAATACATAGTCTTTC 78840	

LOCUS	AC026541/C	RESULT	8
DEFINITION	AC026541 172010 bp DNA	HTG	22-MAR-2000
ACCESSION	Homo sapiens chromosome 1 clone RP11-136B23 map 1, WORKING DRAFT SEQUENCE, 24 unordered pieces.		
VERSION	AC026541		
KEYWORDS	AC026541.1 GI:7283245		
SOURCE	HTG; HTGS_PHASE1; HTGS_DRAFT.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 172010)		
JOURNAL	Birren, B., Linton, L., Nusbaum, C. and Lander, E.		
REFERENCE	Homo sapiens chromosome 1, clone RP11-136B23		
AUTHORS	Unpublished		
	2 (bases 1 to 172010)		
	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,		
	Anderson, S., Baldwin, J., Barna, N., Baskien, V., Beda, F.,		
	Boguslavsky, L., Boukngalter, B., Brown, A., Burkett, G.,		
	Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,		
	Collamore, A., Cooke, P., Deatfield, K., Dewar, K., Diaz, J. S.,		
	Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,		
	Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,		
	Grand-Perre, N., Grant, G., Hages, B., Heaford, A., Horton, L.,		
	Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,		
	Klein, J., Lacombe, K., Lamazares, R., Landers, T., Lehoczy, J.,		
	Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,		
	McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,		
	Melidiri, J., Menes, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J		
	Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,		
	O'Neil, D., Olyart, T. M., Oliver, J., Peterson, K., Pierre, N.,		
	Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,		
	Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,		

TITLE
JOURNAL
COMMENT

Stange-Rhoman, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Testa, S., Theodore, J., Tittell, A., Travers, M., Triggillo, J.,
 Vassiliev, H., Vleider, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

 'Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center: MIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

 Project Information
 Center project name: L7147
 Center clone name: 136.B.23

 Summary Statistics
 Sequencing vector: M13; M7815; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 160176 bases at least Q40
 Consensus quality: 166084 bases at least Q30
 Consensus quality: 168396 bases at least Q20
 Insert size: 173000; agarose-*fp*
 Insert size: 169710; sum-of-*contigs*
 Quality coverage: 4.2 in Q20 bases; agarose-*fp*
 Quality coverage: 4.3 in Q20 bases; sum-of-*contigs*

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

193	1922: contlg of 1922 bp in length
1933	2022: gap of 100 bp
2023	3258: contlg of 1226 bp in length
3259	3358: gap of 100 bp
3359	5825: contlg of 2457 bp in length
5826	5925: gap of 100 bp
5926	8229: contlg of 2304 bp in length
8230	8329: gap of 100 bp
8330	11985: contlg of 3656 bp in length
11986	12085: gap of 100 bp
12086	15487: contlg of 3402 bp in length
15488	15587: gap of 100 bp
15588	18884: contlg of 3297 bp in length
18885	18984: gap of 100 bp
18985	22912: contlg of 3928 bp in length
22913	23013: gap of 100 bp
23013	28119: contlg of 5107 bp in length
28120	28219: gap of 100 bp
28220	33161: contlg of 4942 bp in length
33162	33261: gap of 100 bp
33262	33975: contlg of 5714 bp in length
33976	39075: gap of 100 bp
39076	45608: contlg of 6533 bp in length
45609	45708: gap of 100 bp
45709	52861: contlg of 7153 bp in length
52862	52961: gap of 100 bp
52962	60202: contlg of 7241 bp in length
60203	60302: gap of 100 bp
60303	66671: contlg of 8369 bp in length
66672	68771: gap of 100 bp
68772	77157: contlg of 8386 bp in length
77158	77257: gap of 100 bp
77258	85526: contlg of 8269 bp in length
85527	85626: gap of 100 bp
85627	96744: contlg of 11118 bp in length

Db 93710 GAGCCTGGCCGAGAGAGCCCTGAGGTGGTGGGGCAATGTCCTT 93651
Qy 2172 ggggtggtctcttcttcttcaagttatctgaaatctggtggagtgatcccat 2231
Db 93650 GGGGTGTTCTGCTTATCTTCAAGTTTARCTGAATCTGTGGGAGTGAATCCCAT 93591
Qy 2232 gtgcgaataggggtctctgacctgagatcttcccatctcaatcttctctcatgaag 2291
Db 93590 GTTGCATATGAGAGTCTGCGCCTGAGATCTCCCATCTCAAGTTTCCATGAAAG 93531
Qy 2292 agtactgttaatacatatgctc 2315
Db 93530 ACTACCTGTAATACATAGTCTTC 93507
RESULT 9
AC006544 234523 bp DNA PRI 27-MAY-2000
LOCUS Homo sapiens chromosome 2p13 clone bac91a19, complete sequence.
DEFINITION AC006544
AC006544
VERSION AC006544.19 GI:4630756
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Homo sapiens Chromosome 2p13 BAC Clone bac91a19
REFERENCE
AUTHORS 2 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (09-FEB-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 3 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (17-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 4 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (21-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 5 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (24-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 6 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (26-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS 7 (bases 1 to 234523)
TITLE Chen, F., Do, A., Do, T., Weisler, M. and Roe, B. A.
JOURNAL Direct Submission
Submitted (27-MAY-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Apr 21, 1999 this sequence version replaced gi:4587623.
FEATURES
source
1. 234523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2p13"
/clone="bac91a19"

BASE COUNT 64562 a 50176 c 52136 g 67649 t
ORIGIN
Query Match 50.3%; Score 1164; DB 86; Length 234523;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1152 agaagctcagtagtcttgaagttcccgctgagcagctgagcagatgtgcttgcagatg 1211
Db 161773 AGAAGCTCAGTAGTGTGTTGAAGTCCCGGTGCTACAGCTGCCATGTGCTTGGCATG 161832
Qy 1212 ttctcaagctcagcagatggtcctctgtgtgtgacacagccagtgagtgactgact 1271
Db 161833 TTCTCCAGCTCAGCATGGGCAATCCTGTGTGGACACAGCCAGTAGTGAGTGAAGTCACT 161892
Qy 1272 caccctccgagcagatgagcagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1331
Db 161893 CACCCCTCCGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 161952
Qy 1332 ctgtgaatgcaagaggtgtggtgagcagccatgtgtgtgagcagcttgaagtgagtgag 1391
Db 161953 CTGTGAATGTCAGAGAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162012
Qy 1392 cagccagcagatgctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1451
Db 162013 CAGCCAGCAGATGCTCCGAGCGGGCCACACAGTGGGGGCGGAGCTGGCTGGCTCTT 162072
Qy 1452 ctgggagatctcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1511
Db 162073 CTGGGAGATCTCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 162132
Qy 1512 ggcagagaggaactctctgctgagcagcagcagcagcagcagcagcagcagcagcagcagc 1571
Db 162133 GGCAGAGGGAACCTTCTGCTGAGACAGAGTGGGCTGAGCTGGGCTCCACCTTC 162192
Qy 1572 tgggacacacagctacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1631
Db 162193 TGGGACACACAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 162252
Qy 1632 gctggccttgccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1691
Db 162253 GCTGGCCTTGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162312
Qy 1692 tccctggccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1751
Db 162313 TCCTTGCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 162372
Qy 1752 aacatccatctagagctggtgcaaatgacacactagtgatagatgacatgagcagcagcag 1811
Db 162373 AACATCCATCTAGAGCTGGGCAATGATGACACTAGTATGATGATGATGATGATGATGATG 162432
Qy 1812 gaccactgagatgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1871
Db 162433 GACCACCTGAGATCTGGGGGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162492
Qy 1872 ttgagtagagtagtactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1931
Db 162493 TTGAGTAGTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 162552
Qy 1932 gggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1991
Db 162553 GGGCAGAGCTTGGCGTGAATTCCTGATTCATGAGAAATCAGAACTGCTTGTGGAGCA 162612
Qy 1992 aatcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2051
Db 162613 AATCAGAGCTTGGCGTGAATTCCTGATTCATGAGAACTGCTTGTGGAGCA 162672
Qy 2052 tgggacacatgagcttgggagtggtggtggtggtggtggtggtggtggtggtggtggtggtg 2111
Db 162673 TGGGACCATGATTTGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 162732
Qy 2112 gagcctggtccggaaggaagagccctggaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2171

```

|||||
Db 162733 GAGCCTGGCCGAGAGAGAGACCCCTGGAGCTGGTGGGCGCAATGTGCCCTGAGTCTT 162792
Oy 2172 999gtgtctgtcttatttcttcaagtttcttgaatctgtg99gaagtgcagatcccat 2231
Db 162793 GGGGTCGTCCTCTATTATCTTCAAGTTATCTGCAATCTGTGGGAGTGCATGATCCCAT 162852
Oy 2232 gtctgaattgagatctctgccttgagatctcccatctcagtttcttccatgaaga 2291
Db 162853 GTTGCATATGAGAGTCTCTGCCCTGAGATCTTCCCATCTCAGTTTCTTCCATGAAAG 162912
Oy 2292 agtacctgaatactagttctc 2315
Db 162913 AGTACCTGTAATACATAGTCTTC 162936

RESULT 10
AC007387/c
LOCUS AC007387. 163903 bp DNA PRI 21-DEC-1999
DEFINITION Homo sapiens BAC clone RP11-372J12 from 2, complete sequence.
ACCESSION AC007387
VERSION AC007387.3 GI:5931456
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
99063792
MEDLINE
REFERENCE
AUTHORS Cordes, M., Stromolt, C., Mauplin, R., Yoakum, M. and Le, T.
TITLE The sequence of Homo sapiens BAC clone RP11-372J12
JOURNAL Unpublished.
REFERENCE
AUTHORS Waterston, R.H.
TITLE 3 (bases 1 to 163903)
JOURNAL Direct Submission
Submitted (25-APR-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 163903)
JOURNAL Waterston, R.H.
TITLE Direct Submission
Submitted (28-SEP-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 163903)
JOURNAL Waterston, R.
TITLE Direct Submission
Submitted (21-DEC-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 28, 1999 this sequence version replaced gi:5001502.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0372J12

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence

from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanesse, J.J. and de Jong, P.O. (1996) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-140K4, 200 bp overlap. Actual start of this clone is at base position 1 of RP11-372J12. actual end is at base position 35689 of RP11-140K4.

```

FEATURES
SOURCE
1..163903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-372J12"
/clone_lib="RPCT-11"
1..246
/rpt_family="Retroviral"
repeat_region
326..544
/rpt_family="L1"
repeat_region
1679..1721
/rpt_family="L1"
repeat_region
1731..1848
/rpt_family="L2"
repeat_region
2443..2742
/rpt_family="Alu"
repeat_region
2756..2867
/rpt_family="Retroviral"
repeat_region
2868..3166
/rpt_family="Alu"
repeat_region
3167..3206
/rpt_family="Retroviral"
repeat_region
3391..3690
/rpt_family="Alu"
repeat_region
3767..4048
/rpt_family="Alu"
repeat_region
4049..4068
/rpt_family="L1"
repeat_region
4139..4252
/rpt_family="L1"
repeat_region
4452..4689
/rpt_family="MER1_type"
repeat_region
4690..4991
/rpt_family="Alu"
repeat_region
4992..5055
/rpt_family="MER1_type"
repeat_region
5060..5082
/rpt_family="TTTA)n"
repeat_region
5083..5366
/rpt_family="Alu"
repeat_region
5367..6064
/rpt_family="MER1_type"
repeat_region
6131..6526
/rpt_family="MER1_type"
repeat_region
6544..6594
/rpt_family="MER1_type"
repeat_region
7336..7519
/rpt_family="MER1_type"

```

```
repeat_region      8119..8395  
/rpt_family="Alu"  
repeat_region      8429..9073  
/rpt_family="L1"  
repeat_region      9068..10788  
/rpt_family="L1"  
repeat_region      10796..10913  
/rpt_family="L1"  
repeat_region      11056..11468  
/rpt_family="MaLR"  
repeat_region      11544..11651  
/rpt_family="L2"  
repeat_region      11652..12121  
/rpt_family="MaLR"  
repeat_region      12122..12395  
/rpt_family="L2"  
repeat_region      12396..12700  
/rpt_family="Alu"  
repeat_region      12709..13002  
/rpt_family="Alu"  
repeat_region      13003..13135  
/rpt_family="L2"  
repeat_region      13136..13466  
/rpt_family="Alu"  
repeat_region      13467..13604  
/rpt_family="L2"  
repeat_region      13870..14041  
/rpt_family="MERL_type"  
repeat_region      14282..14416  
/rpt_family="Alu"  
repeat_region      14417..14722  
/rpt_family="Alu"  
repeat_region      14723..14892  
/rpt_family="Alu"  
repeat_region      14894..15559  
/rpt_family="L1"  
repeat_region      15566..16462  
/rpt_family="Retroviral"  
15742..16984  
/note="CPG_Island (%GC=71.8, o/e=0.70, #CPGs=89)"  
16384..17049  
/note="semaphorin Y pseudogene"  
16463..16567  
/rpt_family="Retroviral"  
repeat_region      16568..17529  
/rpt_family="Retroviral"  
repeat_region      17530..17608  
/rpt_family="TAAA)n"  
repeat_region      18007..18797  
/rpt_family="L1"  
repeat_region      18798..18858  
/rpt_family="TAAA)n"  
repeat_region      18944..18969  
/rpt_family="(TTTG)n"  
repeat_region      19031..19254  
/rpt_family="Alu"  
repeat_region      19496..19612  
/rpt_family="L1"  
repeat_region      19689..19993  
/rpt_family="Alu"  
repeat_region      20471..20599  
/rpt_family="Retroviral"  
repeat_region      20668..20978  
/rpt_family="Alu"  
repeat_region      20979..21250  
/rpt_family="Alu"  
repeat_region      21270..21507  
/rpt_family="T2_type"  
repeat_region      21536..21862  
/rpt_family="MERL_type"  
repeat_region      21867..22169  
/rpt_family="Alu"
```

```
Query Match      45.1%; Score 1043; DB 86; Length 163903;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1163; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
  
QY 1152 agaaagtcagtagtgcgtttaaagttccggtgctaaagctgcgcagtgtgtcttgcacg 1211  
|||||  
DB 133048 AGAACGTCAGTAGTGTGTTAAGTTCCTCGCTACAGCTGCGCATGTGCTTGCATG 132989  
|||||  
QY 1212 ttctcaagctcagcatgggcatcctgtgtgtgaccacacacagtcagtcagtcagtc 1271  
|||||  
DB 132988 TTCTCCAACTCAGCATGGCATCTCTGTGTGTGACACACGCCAGTGTGAGTGTGACT 132929  
|||||  
QY 1272 caaccccccgcggagatgagctgagtgagtggtgagaccacaggggccaatgggcctatgc 1331  
|||||  
DB 132928 CACCCCGCGGGGATGACTGTGAGTGTGTGTGACCCAGGGGCCATGGGGCTTATGC 132869  
|||||  
QY 1332 ctgtgaatgtcaggaaggtgtgggcagcccatgtgtgtacagcttaacagcttgataggg 1391  
|||||  
DB 132868 CTGTGAATGTCAGAGAGGTGGGGCAGCCCATGTGTGACAGCTTACAGCTGTGATGGGG 132809  
|||||  
QY 1392 caqccacagagatgctccgcagacggggcccaacagtggtgggggagtcgtgctctt 1451  
|||||  
DB 132808 CAGCCACGAGATGCTCCGAGCCCGGCCACAGTGGGGGGGAGACTGGCTGCTTCTT 132749  
|||||  
QY 1452 ctggggatctcgcagacatccctgacatcctatctgattgtgtgcgtcaagcagcag 1511  
|||||  
DB 132748 CTGGGGGATTCGCGACATCCCGACTCATCTGTGATGTGTGTGGGCTCAGCAGCGCAGC 132689  
|||||  
QY 1512 gcgaacagaggaactctgctctagaagcaaggtgtggcctgagctgggggtcccacttc 1571  
|||||  
DB 132688 GCGACAGAGGGAACTTGTGCTAGAGACAAAGTGGGCTGAGACTGGGGGCTCCACCTTC 132629  
|||||  
QY 1572 tgggacacaaagctacagccaagaacccctccctccctccctccctccctccctccctcc 1631  
|||||  
DB 132628 TGGGACCAACAAGCTACAGCCCAAGACCTCCCTCCCTCTCTCTGAAATAGCGGTTGCC 132569  
|||||  
QY 1632 gctggcccttgccaagagagggcagtggtcttggttgatcttcacacacccctcctgctga 1691  
|||||  
DB 132568 GCT-GCCCTGGCCAAAGAGGGGCACTGGCTTGGTGGATTCACACCACTTCCTGCTTGA 132510  
|||||  
QY 1692 tccttgcccaagccagccacatctgcttaactgggtggtccctcctagccacatgtatga 1751  
|||||  
DB 132509 TCCTTGGCCCAAGGCCACCCACATTCGGCTAACCTGGGGGCTCCTTACCCACATGTGATGA 132450  
|||||  
QY 1752 aacatcatctagagctgggccaatgacacactagtgatgataagatgataacacggagagtc 1811  
|||||  
DB 132449 AACATCATCTAGAGCTGGGCAATGACACTAGTGTATATGATGATACGTGAACGGAGT 132390  
|||||  
QY 1812 gaccactgaagatgctgggggtcactgggcttggaagaacacatcccaagcctctgagtlctct 1871  
|||||  
DB 132389 GACCACCTGAGATGCTGGGGGTCACTGGGCTGGAAGACATCCAGCCTGTGAGTTCTCT 132330  
|||||  
QY 1872 ttgaatgagtgatctacttggaatttagtatctgtctctctctgagcctgagatggcttg 1931  
|||||  
DB 132329 TTGAGTATGAGTATCTTGATTTAGTATCTGTCTCTGTGAGCCTGTGATGGGCTTG 132270  
|||||  
QY 1932 gggccaaaccttgccgcatctcctgattcccatgaaagaacaaacgctctcttgagca 1991  
|||||  
DB 132269 GGGCCAGACCTTGTGCTGATTCCTGATTCCTGATGAGAAATGAGACGTCTTCTGCGCA 132210  
|||||  
QY 1992 aatcaaggctcccccataaatctgaactcctgtaaacctcaccctggcccccatact 2051  
|||||  
DB 132209 AATCAGGGCTTCCCCCTTAACATGTGACATCTCTGTAAACCTTCAATCCCTGGCCCCCTATCT 132150  
|||||  
QY 2052 tggggccatagtttggggatgtgggcaagggacataagacttgcttcttggttg 2111  
|||||  
DB 132149 TGGGCCCATTAGTGTGGGATGGGGCAGAGGAGATGATGATGACTTGTCTTGCTTG 132090  
|||||  
QY 2112 gagccctggcggagagagagccctggaggtgtgttggtgggccaatgtggcctgaagtcctt 2171  
|||||  
DB 132089 GAGCCTGGCCGGAAGAGAGCCCTGAGAGTGTGTGGGGGCAAAATGTGCTTGAAGTCTT 132030  
|||||  
QY 2172 ggggtgtctgtcttacttctcaagttatctgaatcgtgtgggagtgatgatacccat 2231
```



```
* 13687 18688: contig of 5002 bp in length
* 18689 18788: gap of 100 bp
* 18789 26165: contig of 7377 bp in length
* 26166 26265: gap of 100 bp
* 26266 52879: contig of 26614 bp in length
* 52880 52979: gap of 100 bp
* 52980 82107: contig of 29128 bp in length
* 82108 82207: gap of 100 bp
* 82208 88054: contig of 5847 bp in length
* 88055 91774: contig of 3620 bp in length
* 91775 91874: gap of 100 bp
* 91875 101902: contig of 10028 bp in length
* 101903 102002: gap of 100 bp
* 102003 109337: contig of 7335 bp in length
* 109338 109437: gap of 100 bp
* 109438 113399: contig of 3962 bp in length
* 113400 113499: gap of 100 bp
* 113500 118719: contig of 5220 bp in length
* 118720 118819: gap of 100 bp
* 118820 123815: contig of 4996 bp in length
* 123816 123915: gap of 100 bp
* 123916 132492: contig of 8577 bp in length
* 132493 132592: gap of 100 bp
* 132593 137709: contig of 5117 bp in length
* 137710 137809: gap of 100 bp
* 137810 149420: contig of 11611 bp in length
* 149421 149520: gap of 100 bp
* 149521 156195: contig of 6675 bp in length
* 156196 156295: gap of 100 bp
* 156296 171964: contig of 15669 bp in length.
```

FEATURES

source

```
1. .171964
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-136B23"
/clone_1lb="RP11-1.1"
1. .3477
/note="assembly_fragment:00781
clone_end:Sp6
vector_side:left"
misc_feature
3578. .13586
/note="assembly_fragment:00011"
misc_feature
13687. .18688
/note="assembly_fragment:00077"
misc_feature
18789. .26165
/note="assembly_fragment:00151"
misc_feature
26266. .52879
/note="assembly_fragment:00273"
misc_feature
52980. .82107
/note="assembly_fragment:00398"
misc_feature
82208. .88054
/note="assembly_fragment:00431"
misc_feature
88155. .91774
/note="assembly_fragment:00477"
misc_feature
91875. .101902
/note="assembly_fragment:00705"
misc_feature
102003. .109337
/note="assembly_fragment:00744"
misc_feature
109338. .113399
/note="assembly_fragment:00927"
misc_feature
113500. .118719
/note="assembly_fragment:00984"
misc_feature
118820. .123815
/note="assembly_fragment:00995"
misc_feature
123916. .132492
/note="assembly_fragment:01012"
misc_feature
132593. .137709
/note="assembly_fragment:01113"
misc_feature
137810. .149420
/note="assembly_fragment:01399"
misc_feature
149521. .156195
```

```
misc_feature
/note="assembly_fragment:01504"
156296. .171964
/note="assembly_fragment:00674
clone_end:r7
vector_side:right"
BASE COUNT 48050 a 38057 c 35574 g 48570 t 1713 others
ORIGIN
```

```
Query Match 9.7%; Score 225; DB 81; Length 171964;
Best Local Similarity 100.0%; Pred. No.5.2e-118;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 599 agtcaccacacaaacatgaagctccgcgacattggctcctctctcctcgtcgtacc 658
|||||
Db 107851 AGTCATCACCACAAACATGATGAGCTCCGGACCTTGGCTCATCTCTCCCTGCTGACC 107910
|||||
QY 659 ggcgtactaaccttcacatccgcggaccacccactcatgagcaagcgagtlcttcagctatg 718
|||||
Db 107911 GCGTACTACACTTCATCCGGGACCACTCATGATGAGCAGGCGCAGTGTTCACACTGATG 107970
|||||
QY 719 gccacccccctgctgttactacagatagcctatcttcagagtgctgtgtccacaggtgtga 778
|||||
Db 107971 GCCACCCCTGCTGCTGCTACTACATACAGCCTATCTCAGAGTGTGTGCCACAGGCTGA 108030
|||||
QY 779 ccaagccctcagggaaagatgatgtctctcctcctcagggagacag 823
|||||
Db 108031 CCACCTCTCAGGGAAGAGATGATGTGCTCTCTACTGCGGACAG 108075
|||||
```

RESULT 13

AC026541

LOCUS

HOMO SAPIENS CHROMOSOME 1 CLONE RP11-136B23 MAP 1, WORKING DRAFT
SEQUENCE, 24 UNORDERED PIECES.

AC026541

VERSION

AC026541.1

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

1 (bases 1 to 172010)
Homo sapiens chromosome 1, clone RP11-136B23
Unpublished
2 (bases 1 to 172010)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Bouckhalter, B., Brown, A., Buckett, G.,
Campiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., S.,
Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyle, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPeckers, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Menga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Plerre, N.,
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,
Roy, A., Santos, R., Schauer, S., Severly, S., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Triggiani, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

Db	51491	AGTGATCATCCACAAACAAATGAAAGCTCCGGACACTTTGGCTCATCTCTCTCCCTGCTGACCC	51550
0y	659	ggctactactacttcatctccgggaccacccacatctatggagacaggccagtgttctcagctatg	718
Db	51551	GGGTACTCACCCTTCAATCCGGGAGCACCCACATGACATGGACAGGCCAGTGTTCACACTGATG	51610
0y	719	ggcaacccccctgtgctcactacagatacagctatctcagatgcgtggccacaagggtga	778
Db	51611	GGCACCCTCTCTGTGCTACTCTACATATACAGCTTCTTCAGAGTGTGTGGCCACAGGGTGA	51670
0y	779	ccagccctctcaggaggaagatgatgtgcctcctcactccctgggagacag	823
Db	51671	CCAGGCTCTCAGGGGAAGAGTATCATGTGTCCTACCTGCGGACAG	51715
RESULT	14		
LOCUS	G22417	413 bp	DNA
DEFINITION	human STS WI-13990, sequence tagged site.	STS	31-MAY-1996
ACCESSION	G22417		
VERSION	G22417.1	GI:1342743	
KEYWORDS	STS; STS sequence; primer; sequence tagged site.		
SOURCE	human STS derived from sequences in dbest and the Unigene collection.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 413)		
TITLE	Hudson,T.		
JOURNAL	Whitehead Institute/MIT Center for Genome Research; Physically mapped STS		
COMMENT	Unpublished (1995)		
	Contact: Thomas Hudson Whitehead Institute/MIT Center for Genome Research Whitehead Institute for Biomedical Research 9 Cambridge Center, Cambridge MA 02142 USA Tel: 617 252 1900 Fax: 617 252 1902 Email: thudson@genome.wi.mit.edu		
	Primer A: TTACACGTACTCTTCATCGAAGG Primer B: AGTCCCTGGGGTGTTCTG STS size: 138 PCR Profile: Presoak: Denaturation: Annealing: 56 degrees C Polymerization: PCR cycles: 35 Thermal Cycler: Protocol: Template: 10 ng Primer: each 5 pm dNTPs: each 4 mM Tag Polymerase: 0.025 units/ul Total Vol: 20 ul		
	Buffer: MgCl2: 1.5 mM KCl: 50 mM Tris-HCL: 10 mM pH: 9.3		
FEATURES	Derived from dbEST (genbank accession H10410).		
SOURCE	Location/Qualifiers 1. 413 /organism="Homo sapiens" /db_xref="taxon:9606" /map="328.8 CR from top of Chr2 linkage group" 30. 167 30. 153 complement(149..167)		

BASE COUNT	112 a	100 c	102 g	94 t	5 others			
ORIGIN								
Query Match	9.1%; Score 211; DB 54; Length 413;							
Best Local Similarity	100.0%; Pred. No. 7.7e-110;							
Matches 211;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
OY	2105	ctggttgagcctggcgcggaaagaagacctgtgagtggtgttgagggaacaatgtgcctg	2164					
Db	227	CTGCTTGGAGAGCTGGCCGCGAAGAGAGACCCCTGGAGGTGGTGGGGCGCAATGTCCCTG	168					
OY	2165	agctcttggggtggtgttcgtctatctctcaagttatctgaatctctgtggagatgcaga	2224					
Db	167	AGTCCTTGGGGGGGTGGCTGCTTATTTCTTCAATTATCTGAATCTGTGGGAGTGCATGA	108					
OY	2225	tcacctatgtgcaaatgagagctctgcacctgagatcttcccatctcaagtttctcttc	2284					
Db	107	TCCCATGTTGCAATATGAGAGCTCTGCCCTTGACATCTTCCCATCTCAGTTTCTTCC	48					
OY	2285	atgaaagagtagctgtlaatacatagttc	2315					
Db	47	ATGAAGAAGTAGCTGTAATACATAGTGTTC	17					
RESULT	15							
AB047604								
LOCUS	AB047604	2669 bp	mRNA	PRI	01-SEP-2000			
DEFINITION	Macaca fascicularis brain cDNA, clone:QnpA-10727.							
ACCESSION	AB047604							
VERSION	AB047604.1	GI:9929942						
KEYWORDS	fls (full insert sequence).							
SOURCE	Macaca fascicularis adult male brain parietal lobe cDNA to mRNA, clone, fibromacaque brain cDNA library QnpA clone:QnpA-10727.							
ORGANISM	Macaca fascicularis							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;							
	Cercopithecoinae; Macaca.							
	1 (sites)							
AUTHORS	Osada, N., Suzuki, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,							
	Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.							
TITLE	Isolation of full-length cDNA clones from macaque brain cDNA libraries							
JOURNAL	Unpublished (2000)							
REFERENCE	2 (bases 1 to 2669)							
AUTHORS	Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.							
TITLE	Direct Submission							
JOURNAL	Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases.							
COMMENT	Katsuyuki Hashimoto, National Institute of Infectious Diseases,							
	Division of Genetic Resources, 23-1, Toyama 1-chome, Shinjuku-ku,							
	Tokyo 162-8640, Japan (E-mail:khashi@nih.go.jp).							
	URL: http://www.nih.go.jp/yoken/genbank/							
FEATURES	Tel: 81-3-5285-1111(ex.2120), Fax: 81-3-5285-1181)							
	Lab host: TOP10							
	Vector: pME18S-FL3 (Acc.No. AB009864)							
	R. Site1: DraIII (CACGTGTC)							
SOURCE	R. Site2: DraIII (CACCATGTG)							
	Description: 1st strand cDNA was primed with an oligo(dT) primer							
	[ATGTGCTTTTATTTTATTTT]; double-stranded cDNA was synthesized							
	using specific 5' and 3' primers and amplified by PCR. The PCR							
FEATURES	product was digested with SfiI and size selection was performed to							
	exclude fragments <1.5kb. The SfiI-digested PCR product was cloned							
	into distinct DraIII sites of pME18S-FL3. XhoI sites just outside							
	the DraIII sites can be used to isolate the cDNA insert. Libraries							
SOURCE	were constructed by Sugano et al. (University of Tokyo, Institute of							
	Medical Science). Custom primer used for sequencing							
	[5' end primer [CTTCTGCTTCAAAAGCTGG];							
	3' end primer [GACCTCGAGCTCGACGACA]).							

CDS
/clone_lib="macaque brain cDNA library Qmpa"
/dev_stage="adult"
/sex="male"
/tissue_type="brain parietal lobe"
348..794
/codon_start=1
/product="hypothetical protein"
/protein_id="BAB12128.1"
/db_xref="GI:9929943"
/translation="MGAYACBCEQEGGARVVAAYSLVWGSDRDPASRAHTVAGIAGF
FLGVLAASLTLLIGRROORRRORLLARDKVLGLDLPAGPSGTTSYSDPPSPPEDE
RLPLALAKRSGSFGSFPFLDPCPPAHIRLRTGAPLATCDETSL"

BASE COUNT 587 a 669 c 710 g 703 t
ORIGIN

Query Match 6.9%: Score 159; DB 85; Length 2669;
Best Local Similarity 99.2%: Pred No. 8.2e-80;
Matches 259; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1550 tggacctggggctccacctcttgagaccacaagctacagccaagaccctccctccct 1609
DB 580 tggacctggggctccacctcttgagaccacaagctacagccaagaccctccctccct 639
QY 1610 ctccctgaagaatgagcggttgccgctggccctggccaagaaggagtggtcttgatgga 1669
DB 640 ctccctgaagaatgagcggttgccgctggccctggccaagaaggagtggtcttgatgga 699
QY 1670 tctcaccaacctctctgcttgatctcttgccaagcccaagcccaattgggttaactgggg 1729
DB 700 tctcttcacacctctctctgcttgatctcttgccaagcccaagcccaattgggttaactgggg 759
QY 1730 ctcccttagccacatgtgataaacaatccatctagagctgggcaaatgaacactagtgta 1789
DB 760 ctcccttagccacatgtgataaacaatccatctagagctgggcaaatgaacactagtgta 819
QY 1790 taagtgatcaactgggaacggag 1810
DB 820 taagtgatcaactgggaacggag 840

Search completed: May 12, 2001, 18:39:43
Job time: 38184 sec

THIS PAGE BLANK (USPTO)

1. 407

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1877444"
/clone_lib="Soares-NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not I; Site: 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NBHPU, and fetal heart NBH19M) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 118 a 106 c 96 g 87 t

Query Match 14.0%; Score 325; DB 18; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.5e-164;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1991 aatcagagctcccccacacatctgaactcgttaacctatccctgccccctatc 2050
Db 335 AATCAGGGCTTCCCTTAACATCTGACTGTAACCTTATCCCTGCCCCCTATC 276
Qy 2051 ttggagccatgatttgggagatgggacagagagcatagctatgcttctggtt 2110
Db 275 TTGGGCCCATTAAGTTTGGGAGTGGGACACAGGCGCATAGCATGCTTCTTGCTT 216
Qy 2111 ggaagcctggccggaagaagagccctggaggggtgttgggggacaaatgctccggaagctt 2170
Db 215 GGAGCCTGGCCGGAAGAGAGCCCTGGAGGTGGTGGGGCAAAATGTGCCCTGAGTCTT 156
Qy 2171 tggagtgctcgtctcattcctcaagttatctgaactcgttgggagtgatcatcccca 2230
Db 155 TGGGCTGGTCTTCTTATCTTCAATGTTATCTGATCTGTGGGAGTGGATGATGCCCA 96
Qy 2231 tgttgaatatgagatcctcgtccctgagatcctcccatcagtttccctccatgaa 2290
Db 95 TGTTCGAATATGAGTCTGTGCGCTGAGATCTTCCCATCTCATCTTCATCATGAA 36
Qy 2291 gaggatgtaataatagatgttc 2315
Db 35 GAGTACGTGAATACATAGTGTTC 11

RESULT 8
LOCUS H24181 437 bp mRNA EST 06-JUL-1995
DEFINITION ym55f02.r1 Soares Infant Brain INTB Homo sapiens cDNA clone
IMAGE:52280 5' similar to SP:A49069 A49069 COLLAPSEN - ; mRNA
sequence.
ACCESSION H24181
VERSION H24181.1 GI:892876
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 437)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maria, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P., and Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 1989
High quality sequence stops: 310 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.jnl.gov) for further information.
Insert Length: 1989 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 310.

FEATURES

source
Location/Qualifiers
1. 437

/organism="Homo sapiens"
/db_xref="Gene:425216"
/db_xref="taxon:9606"
/clone="IMAGE:52280"
/clone_lib="Soares Infant Brain INTB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda B; Site: 1: Not I; Site: 2: Hind III; 1st strand cDNA was primed with a Not I - Oligo(dT) primer [5', AACTGGAAGATTCGGCGCGGAGGATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda B vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Patricia Bonaldo."

BASE COUNT 105 a 122 c 112 g 95 t 3 others

Query Match 12.9%; Score 298; DB 153; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.3e-149;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 488 gaccacaagacatcggacagtgctgaatgtctcctcgaagaaactaaacatgactga 547
Db 1 GACCACAAAGACATTCGAGACAGTCTGATGTCTCTTCAAGAACTAAACATGACTGCA 60
Qy 548 acaagagagctgctgtcgtgagaaatgattgcccacagcccaactctggagatgactca 607
Db 61 ACAGAGAGACTGCTGTCTGTGACAAATGATGTGCCCAAGCCCAAGCTGAGATGATCA 120
Qy 608 ccaacaacatgagagctcggcagcttggctcattctctcctcggctgagcgtactca 667
Db 121 CCACACATGAAAGATTCGGGACTTGGCTCATCTCTCTCGCTGACCGGCTACTCA 180
Qy 668 ccttaccgagacacacacacatcagagcagcagtgattccagctgtagtgcaccccc 727
Db 181 CCTTATCCGGGAGACACCCACTATGAGAGGACAGTGTTCAGCTGATGGCCACCCCC 240
Qy 728 tgcctgctactacagataagcctatctcagagatcgttggccacaggggtgacagct 785
Db 241 TCGTGTCTACTACAGATACAGAGCTATCTCAGAGTCTGTGCGCCACAGGGTGCACGCT 298

RESULT 9
LOCUS BE734978 1045 bp mRNA EST 15-SEP-2000
DEFINITION 601570573.F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844914 5',
mRNA sequence.
ACCESSION BE734978
VERSION BE734978.1 GI:10149073
KEYWORDS EST.
SOURCE human
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1045)
Mammalia: Eutheria; Primates; Catarrhini; Homidae; Homo.

THIS PAGE BLANK (USPTO)

PA	(SMIK) SWITKLINE BEECHAM PLC.
XX	
PI	Doe TR, Hayes PD, Michalovich D:
XX	
DR	WPI: 1999-407151/35.
XX	P-PSDB: Y23875.
PT	New semaphorin family polypeptide useful for treating spinal injury
XX	and muscular dystrophy
PS	Disclosure: Page 16; 29pp; English.
CC	
CC	The present sequence represents an expressed sequence tag (EST)
CC	sequence, from which DNA encoding SBSEMN1, a semaphorin family
CC	polypeptide, is derived. The polypeptide may be used to screen
CC	for agonists or antagonists including antibodies. Measurement
CC	of the level of SBSEMN1 protein and detection of a mutation in
CC	its polynucleotide may also be used to diagnose a disease or
CC	condition or susceptibility to a disease related to altered
CC	expression or activity of SBSEMN1. These diseases and
CC	conditions include neurodegeneration, spinal injury, neuropathies,
CC	neuromuscular disorders, muscular dystrophy, psychiatric disorders,
CC	inflammatory disorders, developmental malformations, disorders of the
CC	immune system, cancer and viral infections. The protein may also be
CC	used to treat these diseases and conditions by administration as a
CC	vaccine.
SQ	
XX	Sequence 799 BP; 174 A; 237 C; 216 G; 171 T; 1 other;
Query Match	1.3%; Score 54; DB 20; Length 799;
Best Local Similarity	100.0%; Pred. No. 5.1e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	1383 tacagccatctcgaagtcgtgcccacagaaggtagccagcctctcaaggaaaga 1436
Db	258 tacagccatctcgaagtcgtgcccacagaaggtagccagcctctcaaggaaaga 311
RESULT 9	
Q56609/c	
ID	Q56609 standard; DNA: 3821 BP.
XX	
AC	Q56609;
XX	
DT	07-JUL-1994 (first entry)
XX	
DE	Human PCF-2/NT-3 5' flanking region.
XX	
KM	Nerve growth factor-2; NGF-2; neurotrophic factor-3; NT-3;
KW	NGF-2/NT-3; promoter; probe; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	exon
FT	/Location/Qualifiers
FT	1..3506
FT	/tag= a
FT	/codon_start= 3276..3279
FT	IntIntron
FT	3507..3279
FT	/tag= b
FT	CDS
FT	3276..3821
FT	/tag= c
FT	TATA_signal
FT	3244..3249
FT	/tag= d
XX	
PN	EP582796-A.
XX	
PD	16-FEB-1994.
XX	
PF	28-MAY-1993;
XX	
PR	92JP-01040706.
XX	
PR	92JP-0288717.
XX	

[illegible]

THIS PAGE BLANK (USPTO)

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1b="Scars_NhNHPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"

```

NbH4V, and fetal heart NbH4HIV) were mixed, and ss circles were made *in vitro*. Following BAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260323-265523, 480049-3245470, and 484648-488470.

BASE COUNT	118 a	106 c	96 g	87 t
ORIGIN				

Query Match	14.0%	Score 325;	DB 18;	Length 407;
Best Local Similarity	100.0%	Pred. No. 8.5e-164;		
Matches 325; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

[illegible]

Oy 2111 ggaagcttgccggaagaagaccccgagatggttgggggcaaatgtagccttgatctc 2170
Db 215 GGAAGCTTGGCCGGAAGAAGACCCCTGGAGAGTGTTGGGGCAAAATGTGCCCTTGAGTCTT 156

OY 2171 tggsgtgcgtccgcttatctctcgaagttatctcgaaactgtgysgagtgaatgatccccca 2280
DB 155 TGGGCTGGTTTCAGCTTATCTTCAAGTTTATCTGAACCTGTGGGAGTGCGATGCCCA 96			

Oy 2231 tgcgcacataggagctcctcgcccgagatctcccaatcagttctccctcgaatgaaa 2230
 |||||
 Db 95 tgcgcacatatggagctcctcgcccgagatctcccaatcagttctccctcgaatgaaa 36

```
Oy      2291 gagtacgtgtaataacatagtgctc 2315 .
          |||
Db       35 GAGTACGTGTAATAATACATAGTGCTTC 11
```

RESULT	8
H24181	
LOCUS	
DEFINITION	H24181 437 bp mRNA EST 06-JUL-1995
	Y055102.F1 Soares infant brain INTB Homo sapiens CDNA clone
	IMAGE:52280 5' similar to SP:A49069 A49069 COL1A5P1 - ; mRNA
	sequence.

ACCESSION	H24181
VERSION	H24181.1
KEYWORDS	GI:892876
SOURCE	EST.
	human.

ORGANISM	REFERENCE
Homo sapiens	Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holt
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi	1 (passes 1 to 437)
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo	

TITLE
JOURNAL
COMMENT

Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waters
. R., Williamson, A., Wohlmann, P. and Wilson, R.
The Washu-Merck EST Project
Unpublished (1995)
Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Insert Size: 1989
High quality sequence stops: 310 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1989 Std Error: 0.00
Seq primer: M13rpl
High quality sequence stop: 310.
Location/Qualifiers
1..437
/organism="Homo sapiens"
source

```

/db.xref="GDB:425216"
/db.xref="taxon:9606"
/clone="IMAGE:52280"
/clone.lib="Soares Infant brain INIB"
/sex="female"
/dev.stage="73 days post natal"
/lab.host="DH10B (ampicillin resistant)"
/note="Organ: Whole brain; Vector: lambda BD; Site_1: Site_2: Hind III; 1st strand cDNA was primed with a 1- oligo(dT) primer [5', AACTGAGAGAAATTCGCGCCGACGAGAAATTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lambda BD vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fátima Bonaldo."
```

```

Query Match      12.9%; Score 298; DB 153; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.3e-119;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps

```

Qy 488 gaccacaagacatctgcagatgctgtaatggtcccttcagaaactaaacatgactcyca 547
|||||
Db 1 GACCACAGACATTCGGACAGTCTGAATGGTCCCTTCAGAGAACTAAACATGACTGCA 60

DQ 548 acagagacgcctcgtcgtagaacatgatgccccagcccgacctggagatgcataca 607
|||||
|||
Db 61 ACAGAGACGCGCTTCGTGGACAATGATGTCGCCAGCCAGACTGGAGATGCATCA 120

QY 608 ccacacacatgaaagctccggcactctgggcacatctctccctgcctgacccgacgtactca 667
 |||||
 121 ccacacacatgaaagctccggcactctgggcacatctctccctgcctgacccgacgtactca 180

668 ccttcacgcggagaccaccatctgacagcagtgcttcacgtgattgacacccc 727
 |||||
 181 ccttcacgcggagaccaccatctgacagcagtgcttcacgtgattgacacccc 240

QY 728 tgcgtgctactaacagatcacagcctatctcagagtcgtggccacaggtgacacagcct 785
|||||
241 TGCCTGCTACTAACGATCACAGCCTATCTCAGAGTCTGGCCACAGGTGACACAGCCT 298

RESULT	9				
BE734978					
BE734978		1045	hm	mpna	EST
					15-SEP-2000

LOCUS	601570573f1	NT_MCC_21	Homo sapiens	CDNA clone	IMAGE:3844914	5
DEFINITION	mRNA sequence.					
ACCESSION	BE734978					
VERSION	BE734978.1	GI:10149073				

SOURCE	ORGANISM
human.	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	
1 (bases 1 to 1045)	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2001, 10:33:52 ; Search time 454.55 Seconds

(without alignments)
5147.499 Million cell updates/sec

Title: US-09-284-180-1

Perfect score: 4008
Sequence: 1 gccgagggccgcgcagtagc.....aaaaaaaaaaaaaaaa 4008

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 678276 seqs, 291890651 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_0401.*
1: /cgnl_8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /cgnl_8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /cgnl_8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /cgnl_8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /cgnl_8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /cgnl_8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /cgnl_8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /cgnl_8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /cgnl_8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /cgnl_8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /cgnl_8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /cgnl_8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /cgnl_8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /cgnl_8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /cgnl_8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /cgnl_8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /cgnl_8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /cgnl_8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /cgnl_8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /cgnl_8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /cgnl_8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /cgnl_8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4008	100.0	4008	19	Rat semaphorin W e
2	2331	58.2	2331	19	Mouse semaphorin W e
3	107	2.7	4024	21	Human semaphorin W e
4	75	1.9	1761	19	Human semaphorin W e
5	75	1.9	2315	19	Human semaphorin W e
6	75	1.9	2893	20	Human semaphorin W e
7	75	1.9	2894	20	Human semaphorin W e
8	54	1.3	799	20	EST sequence for D
9	48	1.2	3821	15	EST sequence for D
10	48	1.2	3821	20	Human PGF-2/NP-3 5
11	48	1.2	11832	19	Human nerve growth Nucleotide sequenc

C	12	47	1.2	160	19	V16145	Microsatellite mar
C	13	47	1.2	338	20	Z20475	Barley microsatell
C	14	46	1.1	73	13	O34140	Sequence of a micr
C	15	46	1.1	177	20	X61217	DOA1-DDB1 target r
C	16	46	1.1	261	18	T65656	Polymorphic repeat
C	17	46	1.1	300	21	A35342	Mytilus microsat
C	18	46	1.1	367	20	Z27707	Human DNA marker C
C	19	46	1.1	1131	20	Z28294	Rat neuronal limed
C	20	46	1.1	1344	20	X97975	Human secreted pro
C	21	46	1.1	2385	21	A39981	Human TANGO 187 c
C	22	46	1.1	2403	21	A39953	Human TANGO 187-1/
C	23	46	1.1	2418	21	A39979	Human TANGO 187-2
C	24	46	1.1	2446	16	T08768	Rat biglycan cDNA.
C	25	46	1.1	2490	21	A39975	Human TANGO 187-1/
C	26	46	1.1	2523	21	A39978	Human TANGO 187-1/
C	27	46	1.1	2562	21	A39980	Human TANGO 187-3
C	28	46	1.1	2595	21	A39976	Human TANGO 187-2/
C	29	46	1.1	2614	19	V65766	Rat progression el
C	30	46	1.1	2614	20	Z23030	Rat progression el
C	31	46	1.1	2700	21	A39977	Human TANGO 187-1/
C	32	46	1.1	3120	14	O39261	Human t-complex as
C	33	46	1.1	3358	19	V58005	Rat neuritin rece
C	34	46	1.1	3358	21	Z91457	Rat neuritin rece
C	35	46	1.1	3809	12	Q12760	P40 genomic DNA.
C	36	46	1.1	35828	21	A29063	Murine TGF-beta b1
C	37	46	1.1	44576	21	T61522	Cosmid CVO14 conta
C	38	46	1.1	44848	21	A75080	Nucleotide sequenc
C	39	46	1.1	45546	20	T23520	Human kidney amino
C	40	45	1.1	175	18	T65655	Repeat sequence fr
C	41	45	1.1	214	18	T65658	Polymorphic repeat
C	42	45	1.1	241	20	Z20471	Barley microsatell
C	43	45	1.1	260	20	Z20469	Barley microsatell
C	44	45	1.1	262	21	A43685	Human secreted exp

ALIGNMENTS

RESULT	1	
V07279	standard: cDNA to mRNA: 4008 BP.	
AC	V07279;	
DT	08-SEP-1998 (first entry)	
DE	Rat semaphorin W encoding cDNA with 5'UTR and 3'UTR.	
XX		
KW	Rat; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;	
KM	immunosuppressant; gene therapy; diagnosis; research reagent; ds.	
XX		
OS	Rattus norvegicus.	
XX		
FH	Key.	
FT	5'UTR	Location/Qualifiers
FT		1..75
FT		/*tag= a
FT	CDS	76..2406
FT		/*tag= b
FT		/*product= "semaphorin W"
FT	5'UTR	2407..3977
FT		/*tag= c
FT		3978..4008
FT		/*tag= d
PN	W09815628-A1.	
XX		
PD	16-APR-1998.	
XX		
PF	03-OCT-1997;	97MO-JP03549.
XX		
PR	09-OCT-1996;	96JP-0287636.
XX		

[illegible][illegible]


```

XX Human semaphorin W encoding cDNA.
DE
XX Human; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
XX Immunosuppressant; gene therapy; diagnosis; research reagent; ds.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
XX CDS 1..1764
XX FT /*tag= a
XX FT /product= "semaphorin W"
XX
XX MO9815628-A1.
XX
XX 16-APR-1998.
XX
XX 03-OCT-1997; 97WO-JP03549.
XX
XX 09-OCT-1996; 96JP-0287636.
XX
XX (SUMU ) SUMITOMO PHARM CO LTD.
XX
XX Kikuchi K, Kimura T;
XX WPI: 1998-261015/23.
XX P-PSDB: W51314.
XX
XX Nerve extension inhibitor protein semaphorin W - is useful as
XX therapeutic drug and diagnostic and research reagent
XX
XX Claim 2; Page 67-68; 90pp; Japanese.
XX
XX The present sequence encodes human semaphorin W. Semaphorin W and
XX its derivatives are nerve extension inhibitors which are useful as
XX antiallergic, immunosuppressant and anticancer agents. The DNA
XX encoding semaphorin W can also be used in gene therapy, e.g. using
XX a viral vector. The proteins, peptides, DNA and antibodies which
XX recognise the protein or peptides, can be used as diagnostic or
XX research reagents. Semaphorin W can be used as a screen for
XX semaphorin W antagonists with possible therapeutic use.
XX
XX Sequence 1761 BP; 349 A; 511 C; 521 G; 380 T; 0 other.
XX
XX
XX Query Match 1.9%; Score 75; DB 19; Length 1761;
XX Best Local Similarity 100.0%; Pred. No. 8.1e-19;
XX Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 1362 cccctgtgtgctactacagcctatctcagagtcgtgcccacaggtgaccag 1421
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 723 cccctgtgtgctactacagcctatctcagagtcgtgcccacaggtgaccag 782
XX
XX Oy 1422 cctctcagggaaga 1436
XX ||||||||||||||
XX Db 783 cctctcagggaaga 797
XX
XX
XX RESULT 5
XX V07281
XX ID V07281 standard; cDNA to mRNA; 2315 BP.
XX
XX AC V07281;
XX
XX 08-SEP-1998 (first entry)
XX
XX Human semaphorin W encoding cDNA with 3'UTR.
XX
XX Human; semaphorin W; nerve extension inhibitor; antiallergic; anticancer;
XX Immunosuppressant; gene therapy; diagnosis; research reagent; ds.
XX Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT CDS 1..1764
FT FT /*tag= a
FT FT /product= "semaphorin W"
FT 5'UTR 1765..2315
FT FT /*tag= b
XX
XX MO9815628-A1.
XX
XX 16-APR-1998.
XX
XX 03-OCT-1997; 97WO-JP03549.
XX
XX 09-OCT-1996; 96JP-0287636.
XX
XX (SUMU ) SUMITOMO PHARM CO LTD.
XX
XX Kikuchi K, Kimura T;
XX WPI: 1998-261015/23.
XX P-PSDB: W51314.
XX
XX Nerve extension inhibitor protein semaphorin W - is useful as
XX therapeutic drug and diagnostic and research reagent
XX
XX Claim 2; Page 65-66; 90pp; Japanese.
XX
XX The present sequence encodes human semaphorin W. Semaphorin W and
XX its derivatives are nerve extension inhibitors which are useful as
XX antiallergic, immunosuppressant and anticancer agents. The DNA
XX encoding semaphorin W can also be used in gene therapy, e.g. using
XX a viral vector. The proteins, peptides, DNA and antibodies which
XX recognise the protein or peptides, can be used as diagnostic or
XX research reagents. Semaphorin W can be used as a screen for
XX semaphorin W antagonists with possible therapeutic use.
XX
XX Sequence 2315 BP; 463 A; 640 C; 670 G; 542 T; 0 other.
XX
XX
XX Query Match 1.9%; Score 75; DB 19; Length 2315;
XX Best Local Similarity 100.0%; Pred. No. 7.7e-19;
XX Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 1362 cccctgtgtgctactacagcctatctcagagtcgtgcccacaggtgaccag 1421
XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 723 cccctgtgtgctactacagcctatctcagagtcgtgcccacaggtgaccag 782
XX
XX Oy 1422 cctctcagggaaga 1436
XX ||||||||||||||
XX Db 783 cctctcagggaaga 797
XX
XX
XX RESULT 6
XX X86126
XX ID X86126 standard; DNA; 2893 BP.
XX
XX AC X86126;
XX
XX 15-SEP-1999 (first entry)
XX
XX DNA encoding SBSEMN1, a semaphorin family polypeptide.
XX
XX SBSEMN1; semaphorin; neurodegeneration; spinal injury; neuropathy;
XX neuromuscular disorder; muscular dystrophy; psychiatric disorder;
XX inflammatory disorder; developmental malformation;
XX immune system disorder; cancer; viral infection; vaccine; ss.
XX Homo sapiens.
XX
XX OS EP933425-A1.
XX
XX PD 04-AUG-1999.
XX

```

PF 30-SEP-1998: 98EP-0203287.
 XX
 PR 30-JUL-1998: 98GB-0016676.
 PR 30-JAN-1998: 98EP-0300693.
 PR 28-JUL-1998: 98GB-0016423.
 XX
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Doe TR, Hayes PD, Michalovich D;
 XX WPI: 1999-407151/35.
 DR P-PSDB: Y23873.
 XX
 PT New semaphorin family polypeptide useful for treating spinal injury
 XX and muscular dystrophy
 PS Claim 7: Page 13-14; 29pp; English.
 XX
 CC The present sequence encodes SBSEMN1, a semaphorin family polypeptide.
 CC The polypeptide may be used to screen for agonists or antagonists
 CC including antibodies. Measurement of the level of SBSEMN1 protein
 CC and detection of a mutation in its polynucleotide may also be used to
 CC diagnose a disease or condition or susceptibility to a disease related
 CC to altered expression or activity of SBSEMN1. These diseases and
 CC conditions include neurodegeneration, spinal injury, neuropathies,
 CC neuromuscular disorders, muscular dystrophy, psychiatric disorders,
 CC inflammatory disorders, developmental malformations, disorders of the
 CC immune system, cancer and viral infections. The protein may also be
 CC used to treat these diseases and conditions by administration as a
 CC vaccine.
 XX
 SQ Sequence 2893 BP; 579 A; 825 C; 842 G; 647 T; 0 other;
 XX
 XX
 Query Match 1.9%; Score 75; DB 20; Length 2893;
 Best Local Similarity 100.0%; Pred. No. 7.4e-19;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1362 cccctcgtgtcactacagatacagcctactcagagtcgtgcccacagggtgaccag 1421
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1282 cccctcgtgtcactacagatacagcctactcagagtcgtgcccacagggtgaccag 1341
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1422 cctctcagggaaga 1436
 ||||||||||||||||
 DB 1342 cctctcagggaaga 1356
 ||||||||||||||||
 RESULT 7
 X86127
 ID X86127 standard; DNA: 2894 BP.
 XX
 AC X86127;
 XX
 DT 15-SEP-1999 (first entry)
 XX
 DE EST sequence for DNA encoding SBSEMN1.
 XX
 KW SBSEMN1; semaphorin; neurodegeneration; spinal injury; neuropathy;
 KW neuromuscular disorder; muscular dystrophy; psychiatric disorder;
 KW inflammatory disorder; developmental malformation; EST;
 KW expressed sequence tag; immune system disorder; cancer;
 KW viral infection; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP933425-A1.
 PD 04-AUG-1999.
 XX
 PF 30-SEP-1998: 98EP-0203287.
 PR 30-JUL-1998: 98GB-0016676.
 PR 30-JAN-1998: 98EP-0300693.
 PR 28-JUL-1998: 98GB-0016423.
 XX

PR 28-JUL-1998: 98GB-0016423.
 XX
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Doe TR, Hayes PD, Michalovich D;
 XX WPI: 1999-407151/35.
 DR P-PSDB: Y23874.
 XX
 PT New semaphorin family polypeptide useful for treating spinal injury
 XX and muscular dystrophy
 PS Claim 18: Page 14-15; 29pp; English.
 XX
 CC The present sequence represents an expressed sequence tag (EST)
 CC sequence, from which DNA encoding SBSEMN1, a semaphorin family
 CC polypeptide, is derived. The polypeptide may be used to screen
 CC for agonists or antagonists including antibodies. Measurement
 CC of the level of SBSEMN1 protein and detection of a mutation in
 CC its polynucleotide may also be used to diagnose a disease or
 CC condition or susceptibility to a disease related to altered
 CC expression or activity of SBSEMN1. These diseases and
 CC conditions include neurodegeneration, spinal injury, neuropathies,
 CC neuromuscular disorders, muscular dystrophy, psychiatric disorders,
 CC inflammatory disorders, developmental malformations, disorders of the
 CC immune system, cancer and viral infections. The protein may also be
 CC used to treat these diseases and conditions by administration as a
 CC vaccine.
 XX
 SQ Sequence 2894 BP; 580 A; 824 C; 841 G; 646 T; 3 other;
 XX
 XX
 Query Match 1.9%; Score 75; DB 20; Length 2894;
 Best Local Similarity 100.0%; Pred. No. 7.4e-19;
 Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1362 cccctcgtgtcactacagatacagcctactcagagtcgtgcccacagggtgaccag 1421
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 1282 cccctcgtgtcactacagatacagcctactcagagtcgtgcccacagggtgaccag 1341
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 OY 1422 cctctcagggaaga 1436
 ||||||||||||||||
 DB 1342 cctctcagggaaga 1356
 ||||||||||||||||
 RESULT 8
 X86128
 ID X86128 standard; DNA: 799 BP.
 XX
 AC X86128;
 XX
 DT 15-SEP-1999 (first entry)
 XX
 DE EST sequence for DNA encoding SBSEMN1.
 XX
 KW SBSEMN1; semaphorin; neurodegeneration; spinal injury; neuropathy;
 KW neuromuscular disorder; muscular dystrophy; psychiatric disorder;
 KW inflammatory disorder; developmental malformation; EST;
 KW expressed sequence tag; immune system disorder; cancer;
 KW viral infection; vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP933425-A1.
 PD 04-AUG-1999.
 XX
 PF 30-SEP-1998: 98EP-0203287.
 PR 30-JUL-1998: 98GB-0016676.
 PR 30-JAN-1998: 98EP-0300693.
 PR 28-JUL-1998: 98GB-0016423.
 XX

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2001, 14:35:12 ; Search time 6014.38 seconds
(without alignments)
3362.619 Million cell updates/sec

Title: us-09-284-180-4

Perfect score: 2315

Sequence: 1 999999tctcctcctatgtctgc.....cgtgtaatacatagtgttc 2315

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 9623517 seqs, 4368049070 residues

Word size : 0

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: gb_estc1:*
2: gb_estc2:*
3: gb_estc3:*
4: gb_estc4:*
5: gb_estc5:*
6: gb_estc6:*
7: gb_estc7:*
8: gb_estc8:*
9: gb_estc9:*
10: gb_estc10:*
11: gb_estc11:*
12: gb_estc12:*
13: gb_estc13:*
14: gb_estc14:*
15: gb_estc15:*
16: gb_estc16:*
17: gb_estc17:*
18: gb_estc18:*
19: gb_estc19:*
20: gb_estc20:*
21: gb_estc21:*
22: gb_estc22:*
23: gb_estc23:*
24: gb_estc24:*
25: gb_estc33:*
26: gb_estc34:*
27: gb_estc35:*
28: gb_estc36:*
29: gb_estc37:*
30: gb_estc38:*
31: gb_estc39:*
32: gb_estc40:*
33: gb_estc41:*
34: gb_estc42:*
35: gb_estc43:*
36: gb_estc44:*
37: gb_estc45:*
38: gb_estc46:*
39: gb_estc47:*
40: gb_estc48:*
41: gb_estc49:*
42: gb_estc50:*
43: gb_estc51:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: em_gss_pln1:*
191: em_gss_pln2:*
192: em_gss_pro:*
193: em_gss_rod1:*
194: em_gss_rod2:*
195: em_gss_rod3:*
196: em_gss_rod4:*
197: em_gss_rod5:*
198: em_gss_vrt1:*
199: em_gss_vrt2:*
200: em_gss_vrt3:*
201: gb_gss1:*
202: gb_gss2:*
203: gb_gss3:*
204: gb_gss4:*
205: gb_gss5:*
206: gb_gss6:*
207: gb_gss7:*
208: gb_gss8:*
209: gb_gss9:*
210: gb_gss10:*
211: gb_gss11:*
212: gb_gss12:*
213: gb_gss13:*
214: gb_gss14:*
215: gb_gss15:*
216: gb_gss16:*
217: gb_gss17:*
218: gb_gss18:*
219: gb_gss19:*
220: gb_gss20:*
221: gb_gss21:*
222: gb_gss22:*
223: gb_gss23:*
224: gb_gss24:*
225: gb_gss25:*
226: gb_gss26:*
227: gb_gss27:*
228: gb_gss28:*
229: gb_gss29:*
230: gb_gss30:*
231: gb_gss31:*
232: gb_gss32:*
233: gb_gss33:*
234: gb_gss34:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	638	27.6	733	122	AW957134
2	628	27.1	733	122	AW953466
3	449	19.4	459	19	AI379133
4	401	17.3	456	104	AI942370
5	375	16.2	503	7	AA459837
6	345	14.9	534	165	BE273314
7	325	14.0	407	18	AI276851
8	298	12.9	437	153	H24181
9	294	12.7	1045	139	BE734978
10	260	11.2	394	153	H24073
11	235	10.2	337	159	Z45329
12	218	9.4	592	205	AO382219
13	217	9.4	364	156	TO9073
14	211	9.1	413	152	H10410
15	195	8.4	317	14	AA937917
16	193	8.3	347	152	H10623
17	193	8.3	517	155	R54387
18	182	7.9	297	159	Z41043

C	19	168	7.3	285	7	AA460024	AA460024	x250g12.s
C	20	157	6.8	558	16	A1133725	A1133725	Hz82355.Hu
C	21	135	5.8	307	155	RS2120	RS2120	y985h07.s1
C	22	117	5.1	233	152	F02243	F02243	HSCOTE022.n
C	23	112	4.8	469	120	AW838037	AW838037	OV3-LT004
C	24	112	4.8	489	117	AW608110	AW608110	OV3-LT004
C	25	112	4.8	489	117	AW608155	AW608155	OV3-LT004
C	26	85	3.7	212	159	Z422773	Z422773	HSCOTE021.n
C	27	71	3.1	230	152	F11335	F11335	HSC2VE071.n
C	28	75	1.5	376	115	AW436751	AW436751	77209.MAR
C	29	33	1.4	232	114	AW346693	AW346693	29273.MAR
C	30	33	1.4	232	114	AW346698	AW346698	29282.MAR
C	31	32	1.4	689	148	BF467265	BF467265	UI-M-CG0p
C	32	30	1.3	522	138	BE667719	BE667719	155505.MA
C	33	30	1.3	556	138	BE683323	BE683323	182687.MA
C	34	24	1.0	208	115	AW436702	AW436702	77141.MAR
C	35	24	1.0	400	201	AQ018630	AQ018630	CIT-HSP-2
C	36	22	1.0	416	147	BF392081	BF392081	UI-R-CAO-
C	37	22	1.0	428	139	BE754071	BE754071	207388.MA
C	38	22	1.0	430	157	T95289	T95289	y643a07.s1
C	39	22	1.0	482	147	BF392768	BF392768	UI-R-CAO-
C	40	22	1.0	804	231	CNS03Jc6	AL246687	Tetradon
C	41	22	1.0	959	231	CNS04R1V	AL303340	Tetradon
C	42	22	1.0	955	231	CNS03JHn	AL246684	Tetradon
C	43	21	0.9	407	225	A2661172	A2661172	1M0539D10
C	44	21	0.9	473	17	A1205835	A1205835	g933t09.x
C	45	21	0.9	475	227	B75268	B75268	PCt11-1406

ALIGNMENTS

Db	11	GAATTTGACCAAGCTGGCTCTGGTTGGCTCCGCTACTGAGGTGACCAAGTAAATAC	70
OY	981	aaccaactgttgccgctctccagaagctgctcaagatgcalctctgagccagaccagctcg	1040
Db	71	AACCAACTGTGGCCGTCTCCAGAGCTGTCACAGTGCATCTCGGCCACGAGACCAGTGTG	130
OY	1041	tgctctgagcttcgcgctctgaatgagtgtgtgtgcccatactgcggggagaccaggaagtgct	1100
Db	131	TGCGTGGAGCTTCGGCTGGATGAGTGTGTGGCCATGCCGGGGAGACCGAGGGTTGGT	190
OY	1101	ccaagacatagagatcacaagaatgtccctctctgtgtctctaaagaagcttgagaagctgccc	1160
Db	191	CCAAAGCATATGATGCACACATATGCTTCCTTTTGTTGTTCTTAAGAGCCTTGGGAACGTCC	250
OY	1161	agtagtcttgaagcttccgcttgactaagaactgtgcacatgtgtctctgtacatgtgtcccaag	1220
Db	251	AGTAGTGTTTGAAGTTCCTCCGTGGCTACAGGTGGCATGTGTGCTTGCCATGTCTCCAAAG	310
OY	1221	ctcagcatgggacatcctgtgtgtgtgcacaagcccagttgaagtacatgcaactcaaccccccg	1280
Db	311	CTCAGCATGGGCATCCGTGTGTGGCACCAAGCCCAATGAGTGAAGTGCACATCACCCCCCG	370
OY	1281	gcggagatggagcttgaaagtgtgtgtgtgagcccccagaggggcatagggcgtctatgtccttgaagt	1340
Db	371	GCGGAGATGAGACTGGAGGTGTGTGTGACCCCAAGGGGSCAATGAGGCGCTTTATGCTGTGAATG	430

RESULT	1
LOCUS	AM957134
DEFINITION	AM957134 733 bp mRNA
ACCESSION	EST369204 MAGC resequences, MAGD Homo sapiens cDNA, mRNA sequence
VERSION	AM957134
KEYWORDS	AM957134.1 GI:8146817
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 733)
	Hogde, P., Ol, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holl, I.E., Seed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL	Unpublished (2000)
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johng@tigr.org Plate: 102
FEATURES	Seq primer: Reverse.
SOURCE	Location/Qualifiers
	1..733
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_lib="MAGC resequences, MAGD"
	/note="Vector: pBluescriptK"
BASE COUNT	134 a 204 c 238 g 157 t
ORIGIN	

Query Match	27.66;	Score 638;	DB 122;	Length 733;
Best local Similarity	99.9%;	Pred. No. 0;		
Matches 688;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps

LOCUS	AW953466	733 bp	EST	01-JUN-2000
DEFINITION	EST346553c MAGE sequences, MAGB Homo sapiens cDNA, mRNA sequence			
ACCESSION	AW953466			
VERSION	AW953466.1	GI:8143149		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
AUTHORS	1 (bases 1 to 733) Hagde,P., Ol,R., Abernathy,K., Dharp,S., Gaspar,R., Gay,C., Holm,I.E., Seed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.			
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johng@tigr.org Plate: 51			

Email: johnq@tigr.org
 Plate: 51

```
FEATURES                               Seq primer: Reverse
source                                Location/Qualifiers
1..733
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE-reserved sequences, MAGB"
/vector="Vector: pbluescriptskm"
BASE COUNT    133 a    207 c    238 g    154 t    1 others
ORIGIN

Query Match      27.1%; Score 628; DB 122; Length 733;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 678; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 921 gaaattgtaccacagctggctcctgtgtgtcctccgtactgtgagtgacaaatgaatc 980
Db 11 GAAATTGTACCAAGCTGGCTCCTGTGTGTGCTCCGTACTGAGTGACAAAGTGAATAC 70

Oy 981 aacacacgtgtgctgtcctcagagctgtcagagtgatcctgtgcccagagacagctc 1040
Db 71 AACCACTGTGTGCTGTCCAGAGCTGTCCAGAGTGATCTGTGCCCAAGACCACTGTG 130

Oy 1041 tgcctggagctcgcggagagtgagtgatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1100
Db 131 TGCCGTGGAGCTTCGGGCTGTGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 190

Oy 1101 ccaagacatagatcagacagatgtcctcctcctcctcctcctcctcctcctcctcctc 1160
Db 191 CCAAGACATAGATCAGACAGATGTCTCCTTGTGTGTCTTAAGAGCCTGGAGAAAGCTGC 250

Oy 1161 agtagtgtttgaagttcccggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1220
Db 251 AGTAGTGTGTGAAGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 310

Oy 1221 ctcagacatgagatcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1280
Db 311 CTCACATGAGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 370

Oy 1281 gctggagatgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1340
Db 371 GCGGATGAGACTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 430

Oy 1341 tcagagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1400
Db 431 TCAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 490

Oy 1401 agatgtcctcagacgcggccacacacacagtggtgtgtgtgtgtgtgtgtgtgtgt 1460
Db 491 AGATGCTCGAGCGCGGCGCCACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 550

Oy 1461 tctcgagcatcctcgtactcctcctcctcctcctcctcctcctcctcctcctcct 1520
Db 551 TCTCGACATCCTCGATCTCTATCTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 610

Oy 1521 ggaacttctgtctagaagcaagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1580
Db 611 GGAATCTTGTGCTTAAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 670

Oy 1581 aagctacagcaagacact 1599
Db 671 AAGCTACAGCAAGACCT 689

RESULT 3
LOCUS AI379133 459 bp mRNA EST 18-MAR-1999
DEFINITION tcs99g09.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:2068960
ACCESSION AI379133
VERSION AI379133.1 GI:4188986
KEYWORDS EST.
SOURCE human.
```

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 459)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1306 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 439.
FEATURES
source
1..459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE:2068960"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below); Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbH, pregnant uterus
NbHPU, and fetal heart NbH19H) were mixed, and 98 clones
were used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT    142 a    117 c    105 g    95 t
ORIGIN

Query Match      19.4%; Score 449; DB 19; Length 459;
Best Local Similarity 100.0%; Pred. No. 9.2e-231;
Matches 449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1867 tctcttgagatgagctgtactgtgatttgaatcgtctcctcctcgtgagctgtgag 1926
Db 459 TCTCTTGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400

Oy 1927 gcttggggcgaacctgtgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1986
Db 399 GCTTGGGGCCAGACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 340

Oy 1987 cagcaaatcagagcttcccccatacatctgaactcctgttaaccttcacctcctgccc 2046
Db 339 CAGCAATCAGAGGCTTCCCTTAACATCTGAACCTCTGTAACCTTCACTCCCTGGCCCC 280

Oy 2047 tatcttggggcccatagtttgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2106
Db 279 TATCTTGGGGCCCATAGTTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 220

Oy 2107 ggttgagacctgtgcccgaagagagccctgtgagtggtgtgtgtgtgtgtgtgtgt 2166
Db 219 GGTTGAGACCTGTGCGGAAGAGCCCTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 160

Oy 2167 tctctggagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2226
Db 159 TCCTTGGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 100

Oy 2227 cccatgttgcataatagagctctgtccctgtgagatcctcccatcctcagtttctcctcat 2286
Db 99 CCCATGTTGCATATAGAGGTCTGTCCCTGTGAGATCTTCCCATTCACATTTCCCTTCAT 40

Oy 2287 gaaagagtaagtgtaataacatagtgctc 2315
```



```
Db 39 GAAAGACTAGCTGAATACATAGTCTTC 11
|||||
RESULT 4
LOCUS AI942370 456 bp mRNA EST 08-MAR-2000
DEFINITION WO80e06.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:246166 3',
ACCESSION AI942370
VERSION AI942370.1 GI:5707026
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 456)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 760 Std Error: 0.00
Seq primer: -40UP from Glibco
High quality sequence stop: 405.
FEATURES
source
1.456
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:246166"
/clone_1lb="NCI_CGAP_Kid11"
/lab_host="DH10B"
/note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid11 was
prepared, and as circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo. "
```

```
BASE COUNT 142 a 116 c 105 g 93 t
ORIGIN
Query Match 17.3%: Score 401; DB 104; Length 456;
Best local similarity 99.8%: Pred. No. 7,7e-205;
Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 1864 agtctctttagatagatgactttagatttagatcttctctctgagctgga 1923
|||||
Db 456 AGTCTCTTGTGATATGATGATTTAGTATCTGTTCTCTCTGAGCCCTGGA 397
|||||
Qy 1924 tgggtctggggccaagacttgcctgattcccatgagaaatcagaactgtt 1993
|||||
Db 396 TGGCCTTGGGGCCAGACCTTTGCCCTGATTCCTCAATGAGAAATCAGAACTGCTTT 337
|||||
Qy 1984 ctgcagcaaatcagagcttcccccataactcgaactccgaactcattccctggcc 2043
|||||
Db 336 CTGACAGAAATCAGAGCTTCCCTTAACATCTGTAACCTCTCATCTCTGACC 277
|||||
Qy 2044 ccctattctggcccatgatttgggattgggacagagcatagctagacttgctt 2103
```

```
Db 276 CCTATCTTGGGCCCATTAATCTTTGGGATGGGACACAGGCGCATAGCATCTTCTTC 217
|||||
Qy 2104 tctgattgagacctgagcgaagaagacctggagtgattggggccaatgctgc 2163
|||||
Db 216 TCTGCTTGGAGAGCTGGCGGAGAGAGAGCCCTGAGAGTGTTGGGGCAAAATGTGCCCT 157
|||||
Qy 2164 gattccttgggtgattctctcttcttctcaagttctcgaatctgtggggaatgcatg 2223
|||||
Db 156 GAGTCCTTGGGGTGTTCTCTCTTATTCTTCAAGTTATCTGAAATCTTGGGGAGTGCATG 97
|||||
Qy 2224 atcccattgtgaatatgagctctgcctcgaatctcccatcagtttcttc 2283
|||||
Db 96 ATCCCATGTTGCAATATGCAATCTGCCCCGAGATCTTCCCATCTCACTTTCTTC 37
|||||
Qy 2284 catgaagagtaactgtaataactagttct 2315
|||||
Db 36 CATGAAGAGTACCTGTAAATACATAGTGTTC 5
|||||
```

```
RESULT 5
LOCUS AA459837 503 bp mRNA EST 09-JUN-1997
DEFINITION 2x30g12.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:795718
ACCESSION AA459837
VERSION AA459837.1 GI:2184744
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 503)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie,
T., Waterston, R., and Wilson, R.
Washu-Merck EST Project 1997
JOURNAL Unpublished (1997)
COMMENT Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 459.
FEATURES
source
1.503
/organism="Homo sapiens"
/db_xref="GDB:6039163"
/db_xref="taxon:9606"
/clone_1lb="IMAGE:795718"
/clone_1lb="Soares_testis_NHT"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - Oligo(dT) primer [5',
TGTTCACATCTGAGTGGAGGCGGCCCATTTTCTTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
```

```
BASE COUNT 105 a 136 c 149 g 113 t
ORIGIN
```

```
Query Match 16.2%: Score 375; DB 7; Length 503;
```


source 1. 407
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:187744"
/clone_lib="Soares-NbHMPu.S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 118 a 106 c 96 g 87 t
ORIGIN

Query Match 14.0%; Score 325; DB 18; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.5e-164;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1991 aaatcaggagctcccccatacatctgaactccctgaacatccatccctggcccccatac 2050
|||||
DB 335 AATATGAGGCTTCCCTTAACATCTGACTCCGTAAACCTTATCCCTGCCCCATC 276
|||||
OY 2051 ttggagcccatagttctggggatggggacagggcatagctatgactttgcttctgct 2110
|||||
DB 275 TTGGGCCCCATTAGTTGGGGATGGGGACAGGCGCATAGCTATGACTTTCTTGCTT 216
|||||
OY 2111 ggaagcctggccggaagagagagcctggagggtggtctgggggcaaatgctccgagctc 2170
|||||
DB 215 GGAGCTGGCCGGAGAGAGAGCCCTGGAGGTGTTGGGGCAAAATGTCCTGAGTCTT 156
|||||
OY 2171 tggggctgctcctcattctcattcaatctcgaatctctgaggagtgatcatcccca 2230
|||||
DB 155 TGGGGGTGTTGCTTATCTTCAAGTTTATCTGATCTGTGGGAGATGATATCCCA 96
|||||
OY 2231 tgttgcataatgagctctcctgagagatctcccatcagtttctcctcatgaaa 2290
|||||
DB 95 TGTTCGAATATGAGTCTGCTGAGATCTGCCCATCTCAGTTTCTTCTTCATGAAA 36
|||||
OY 2291 ggttgcgttgaatacatagtgctc 2315
|||||
DB 35 GAGTACGTGTAATACATAGTGTTC 11
|||||

RESULT 8
LOCUS H24181 437 bp mRNA EST 06-JUL-1995
DEFINITION ym5f02.r1 Soares Infant brain INIB Homo sapiens cDNA clone IMAGE:52280 5' similar to SP:A49069 A49069 COLLA1P5IN - ; mRNA sequence.
H24181
ACCESSION H24181.1 GI:892876
VERSION H24181.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 437)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Woldman, P., and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 1989
High quality sequence stops: 310 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1989 Std Error: 0.00
Seq primer: M13Rpl
High quality sequence stop: 310.

FEATURES
source location/Qualifiers
1. 437
/organism="Homo sapiens"
/db_xref="GDB:425216"
/db_xref="taxon:9606"
/clone="IMAGE:52280"
/clone_lib="Soares Infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lambda BA; Site_1: Not I; Site_2: Hind III; 1st strand cDNA was primed with a Not I - Oligo(dT) primer (5' AACTGGAGATATCGCGCCGAGGAATTTTCTTTTCTTTT 3'); double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 122 c 112 g 95 t 3 others
ORIGIN

Query Match 12.9%; Score 298; DB 153; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.3e-149;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 488 gaccacaagacattcgagagagctgtaattgctcctcagagaaactaaacatgactcga 547
|||||
DB 1 GACCACAGACATTCGGACAGTGTGATGATGCTCCCTTCAAGAACTAATACATGACTGCA 60
|||||
OY 548 acagagagactgctctgctgagacatgaltgccccacagaccagactgagagtgactca 607
|||||
DB 61 ACAGAGAGCTGCTGCTGTGTGAGACATGATGTGCCCCAGCCAGACCTGAGAGTGCATCA 120
|||||
OY 608 ccaacaacatgagagctcggagactttggtctcattctctcctgctgactcagctgactca 667
|||||
DB 121 CCAACAACATGAGAGCTTCGGACATCTTGGCTCATCTCTCCTGCGACCGCTACTCA 180
|||||
OY 668 ccttcatcggagagccaccacatcagagagggcagaggttccagctgagagccacccc 727
|||||
DB 181 CCTTCATCCGGAGACCCACTCATGAGACAGCCAGGTTTCCAGCTGATGGCCACCC 240
|||||
OY 728 tctgtgctacatagatagacatcctatctcagagctgctgagccacagaggtaccagcct 785
|||||
DB 241 TGCCTGCTCATACAGATACAGACCTATCTCAGAGTGTGCGCCACAGGGGTACACAGCTT 298
|||||

RESULT 9
LOCUS BE734978 1045 bp mRNA EST 15-SEP-2000
DEFINITION 601570573f1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3844914 5', mRNA sequence.
BE734978
ACCESSION BE734978
VERSION BE734978.1 GI:10149073
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1045)


```

RESULT 11
245329 337 bp mRNA EST 14-NOV-1994
LOCUS HSCM091 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION c2mg09, mRNA sequence.
ACCESSION 245329
VERSION 245329.1 GI:574541
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 337)
Auffray, C., Behar, G., Bols, F., Bouchier, C., da Silva, C., Devignes
M.D., Duprat, S., Houllatte, R., Jumeau, M., Lamy, B., Lorenzo, F.,
Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,
Sebastiani-Kabakchis, C. and Tessier, A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
JOURNAL Contact: Genethon
MEDLINE Genethon
COMMENT Genethon Centre de recherche sur le Genome Humain
1, rue de l'interationale, BP60 91002 Evry Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C; Genexpress_sequence_id: y1c-2mg09
Seq primer: (-21)M13-universal.
FEATURES
source
1. 337
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2mg09"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid B4; Site_1: HindIII;
Site_2: NotI; sex=Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total brain
total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid B4 vector. Clone library from B.Souares, Psychiatry
Dept. Columbia University, USA. Normalization method:
Bento Soares, P.N.A.S in press"
BASE COUNT 81 a 106 c 82 g 67 t 1 others
ORIGIN
Query Match 10.2%; Score 235; DB 159; Length 337;
Best Local Similarity 99.4%; Pred. No. 3.4e-115;
Matches 335; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 727 ctgtgtgtcactacacagatatactcagagtcgtgagccagaggtgacgctc 786
|||||
DB 241 CTGCTGTACTACTACAGATTACAGGCTATCTCAGAGTGTGGCCACAGAGGTGACAGGCTC 300
OY 787 tcaggaagaagatcatatgtctcttactctgagagacg 823
|||||
DB 301 TCAGGGAACACATATGATGTCTTACTCTGGGACAG 337
RESULT 12
A0352219 592 bp DNA GSS 24-JAN-1999
LOCUS CITR1-El-2536H1.1.TF CITR1-El Homo sapiens genomic clone 2536H1.1,
DEFINITION DNA sequence.
ACCESSION A0352219
VERSION A0352219.1 GI:4179554
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 592)
Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and
Venter, J.C.
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
Unpublished (1997)
JOURNAL Other_GSSs: CITR1-El-2536H1.1.TF
COMMENT Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeet@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
1. 592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2536H1.1"
/clone_lib="CITR1-El"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 140 a 171 c 143 g 136 t 2 others
ORIGIN
Query Match 9.4%; Score 218; DB 205; Length 592;
Best Local Similarity 100.0%; Pred. No. 5.4e-106;
Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


ACCESSION AA937917
VERSION AA937917.1 GI:3096028
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 317)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.

CDNA library preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1661 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 315.
Location/Qualifiers

FEATURES
source 1..317
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1435896"
/clone_1lb="NCI-CGAP_C08"
/issue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaudo."

BASE COUNT 97 a 82 c 77 g 61 t
ORIGIN

Query Match 8 4%; Score 195; DB 14; Length 317;
Best Local Similarity 99.7%; Pred. No. 1.3e-93;
Matches 315; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 2000 ctccccctaacatctgaactccctgaaacctcaccctgccccctatccttgagccca 2059
DB 317 CTTCGCCCTACATCTGAACTCCTGTAAACCTTCATCCCGGCCCTATCTGGGCCCA 258
OY 2060 ctatgttcgggagtgaggcagggatagctatgaacttccttccttggttgagccttg 2119
DB 257 TTAGTTTGGGGATGGGGCACAGGCATGACTTGTCTTCTGGTTGGAGCCTGG 198
OY 2120 cgggaaggaaagcccttgagggtggttgagggaatgctgacctgagctctggggtggt 2179
DB 197 CCGGAAGGAAGACCCCTGAGGTGTTT-GGGGCAATGTCCCTGAGTCTTGGGGGTGT 139
OY 2180 tctgctatctctcaagttatctgaatctgltggggagtgacgacaloccccatgttgcac 2239
DB 138 TCTGCTATTCTTCAAGTTATCTGAATCTGTGGGAGTGCATGATCCCATGTGCAAT 79
OY 2240 atggagctctgcctcgagatcttcccatctcagtttcttcacatgaagaagtagctg 2299
DB 78 ATGAGTCTCTGCCCTGAGATCTTCCCATCTCAGTTTCTTCATGAAGAAGTAGCTG 19
OY 2300 taatacatagtttc 2315
DB 18 TAAATACATAGTCTTC 3

Search completed: May 12, 2001, 14:35:31
Job time: 25224 sec
